A Guidance Manual to Support the Assessment of Contaminated Sediments in Freshwater, Estuarine, and Marine Ecosystems

Volume 3 – Interpretation of the Results of Sediment Quality Investigations

Submitted to:

Mike Macfarlane
British Columbia Ministry of Water, Land, and Air Protection
Pollution Prevention and Remediation Branch
2975 Jutland Road
Victoria, British Columbia
V8V 1X4

Submitted – November 2001 – by:

MacDonald Environmental Sciences Ltd.
#24 - 4800 Island Hwy N.
Nanaimo, British Columbia V9T 1W6

104-07695/

A Guidance Manual to Support the Assessment of Contaminated Sediments in Freshwater, Estuarine, and Marine Ecosystems

Volume 3 – Interpretation of the Results of Sediment Quality Investigations

Submitted to:

Mike Macfarlane
British Columbia Ministry of Water, Land, and Air Protection
Pollution Prevention and Remediation Branch
2975 Jutland Road
Victoria, British Columbia
V8V 1X4

Prepared – November 2001 – by:

Christopher G. Ingersoll[1] **and Donald D. MacDonald**[2]

[1]**United States Geological Survey**
4200 New Haven Road
Columbia, Missouri 65201

[2]**MacDonald Environmental Sciences Ltd.**
#24 - 4800 Island Hwy N.
Nanaimo, British Columbia V9T 1W6

Table of Contents

Table of Contents i
List of Tables iv
List of Figures v
Executive Summary vii
List of Acronyms xi
Glossary of Terms xvi
Acknowledgments xxii

Chapter 1. Introduction 1
1.0 Background 1

Chapter 2. Assessment of Sediment and Pore Water Chemistry 4
2.0 Introduction 4
2.1 Selection of Metrics and Targets for Sediment Chemistry 4
2.2 Availability of Standard Methods 9
2.3 Advantages and Disadvantages of Sediment Chemistry Data 9
2.4 Evaluation of Data Quality 12
2.5 Methodological Uncertainty 13
  2.5.1 Uncertainty Associated with Sediment Chemistry 13
  2.5.2 Uncertainties Associated with Uses of Sediment Quality Guidelines 16
2.6 Interpretation of Data 19
2.7 Recommendations 22

Chapter 3. Sediment Toxicity Testing 25
3.0 Introduction 25
3.1 Selection of Metrics and Targets for Sediment Toxicity 25
3.2 Availability of Standard Methods 29
3.3 Advantages and Disadvantages 33
3.4 Evaluation of Data Quality 34
3.5 Methodological Uncertainty 35
3.6 Interpretation of Data 39
3.7 Recommendations 41

Chapter 4. Benthic Invertebrate Community Assessment 44
4.0 Introduction 44
4.1 Selection of Metrics and Targets for Benthic Invertebrates Community Structure 44
4.2 Availability of Standard Methods 46
4.3 Advantages and Disadvantages 47
4.4 Evaluation of Data Quality 49

4.5 Methodological Uncertainty 51
4.6 Interpretation of Data 53
4.7 Recommendations 54

Chapter 5. Bioaccumulation Assessment 58
5.0 Introduction 58
5.1 Selection of Metrics and Targets for Bioaccumulation Assessment 58
5.2 Availability of Standard Methods 61
5.3 Advantages and Disadvantages 64
5.4 Evaluation of Data Quality 66
5.5 Methodological Uncertainty 67
5.6 Interpretation of Data 70
5.7 Recommendations 74

Chapter 6. Fish Health and Fish Community Assessments 77
6.0 Introduction 77
6.1 Selecting Metrics and Targets in Fisheries Assessments 77
6.2 Availability of Standard Methods 80
6.3 Advantages and Disadvantages 80
6.4 Evaluation of Data Quality 82
6.5 Methodological Uncertainty 83
6.6 Interpretation of Data 84
6.7 Recommendations 85

Chapter 7. Integration of Information on Multiple Indicators of Sediment Quality Conditions 87
7.0 Introduction 87
7.1 Integration of Information on Multiple Indicators of Sediment Quality Conditions 87
  7.1.1 Integration of Information on Multiple Indicators for Assessing Impacts on Sediment-Dwelling Organisms and Other Receptors 90
  7.1.2 Integration of Information on Multiple Indicators of Sediment Quality in the Assessment of Impacts on Wildlife . . . 95
  7.1.3 Integration of Information on Multiple Indicators of Sediment Quality in the Assessment of Impacts on Human Health 97
7.2 Summary 98

Chapter 8. References 99

Appendix 1. Recommended Uses of Sediment Quality Guidelines 117
A1.0 Introduction 117
A1.1 Monitoring Program Design 117
A1.2 Interpretation of Sediment Chemistry Data 118
A1.3 Analysis of Dredged Materials for Open-Water Disposal 122
A1.4 Ecological Risk Assessment 123
A1.5 Development of Sediment Quality Remediation Objectives 125

Appendix 2. Methods for Determining Background Levels of Sediment-Associated Contaminants 127
A2.0 Introduction 127
A2.1 Reference Sediment Approach 128
A2.2 Reference Element Approach 129

Appendix 3. Approaches to the Development of Numerical Sediment Quality Guidelines 131
A3.0 Introduction 131
A3.1 Screening Level Concentration Approach 132
A3.2 Effects Range Approach 133
A3.3 Effects Level Approach 134
A3.4 Apparent Effects Threshold Approach 135
A3.5 Equilibrium Partitioning Approach 136
A3.6 Logistic Regression Modeling Approach 137
A3.7 Consensus Approach 139
A3.8 Tissue Residue Approach 140

Appendix 4. Criteria for Evaluating Candidate Data Sets 143
A4.0 Introduction 143
A4.1 Criteria for Evaluating Whole Sediment, Pore water, and Tissue Chemistry 144
A4.2 Criteria for Evaluating Biological Effects Data 145

List of Tables

Table 2.1 Sediment quality guidelines that reflect threshold effect concentrations (TECs; i.e., below which harmful effects are unlikely to be observed; from MacDonald *et al.* 2000b) 147

Table 2.2 Sediment quality guidelines that reflect probable effect concentrations (PECs; i.e., above which harmful effects are likely to be observed; from MacDonald *et al.* 2000b) 149

Table 2.3 Advantages and disadvantages of whole sediment and pore water chemistry (Ingersoll *et al.* 1997) 151

Table 2.4 Uncertainty associated with sediment chemistry measurements (Ingersoll *et al.* 1997) 152

Table 2.5 Uncertainty associated with sediment quality guidelines (Ingersoll *et al.* 1997) 153

Table 2.6 Summary of potential targets for pore water chemistry 154

Table 3.1 Rating of selection criteria for freshwater sediment toxicity testing organisms (ASTM 2001a; USEPA 2000a) 157

Table 3.2 Summary of standard methods for conducting whole-sediment toxicity or sediment bioaccumulation tests with freshwater invertebrates 158

Table 3.3 Advantages and disadvantages of laboratory sediment toxicity tests (ASTM 2001a; USEPA 2000a) 159

Table 3.4 Test conditions for conducting a 28- to 42-day sediment toxicity test with *Hyalella azteca* (ASTM 2001a; USEPA 2000a) 160

Table 3.5 Test acceptability requirements for a 42-day sediment toxicity test with *Hyalella azteca* (ASTM 2001a; USEPA 2000a) 162

Table 3.6 Uncertainty associated with sediment phases used in laboratory toxicity tests (Ingersoll *et al.* 1997) 164

Table 3.7 Uncertainty associated with endpoints measured in laboratory toxicity tests with sediment (Ingersoll *et al.* 1997) 165

Table 4.1 Uncertainty associated with benthic community assessments (Ingersoll *et al.* 1997) 166

Table 4.2 Advantages and disadvantages of benthic invertebrate community structure data 167

Table 5.1 Selection criteria for sediment bioaccumulation test organisms (ASTM 2001d; USEPA 2000a) 168

Table 5.2 Advantages and disadvantages of tissue chemistry data 169

Table 5.3 Recommended test conditions for conducting a 28-day sediment bioaccumulation test with *Lumbriculus variegatus* (ASTM 2001d; USEPA 2000a) 170

Table 5.4 Test acceptability requirements for a 28-day sediment bioaccumulation test with the oligochaete, *Lumbriculus variegatus* (ASTM 2001d; USEPA 2000a) 172

Table 5.5 Uncertainty associated with bioaccumulation assessments (Ingersoll *et al.* 1997) 174

Table 6.1 Methods for evaluating the effects of exposure to COPCs in fish (from Schmitt *et al.* 2000) 175

Table 6.2 Methodological uncertainty associated with fish health and fish community assessments 177

Table 7.1 Contingency table for assessing impacts of contaminated sediments on aquatic life based on three separate indicators of sediment quality (sediment quality triad adapted from Chapman 1992 and Canfield *et al.* 1996) 178

Table 7.2 Contingency table for assessing impacts of contaminated sediments on aquatic life based on four separate indicators of sediment quality 179

Table 7.3 Contingency table for assessing impacts of contaminated sediments on aquatic life based on two separate indicators of sediment quality 183

Table 7.4 Contingency table for assessing impacts of contaminated sediments on wildlife based on three separate indicators of sediment quality 184

Table 7.5 Contingency table for assessing impacts of contaminated sediments on human health based on two separate indicators of sediment quality 185

Table A1.1 Incidence of toxicity predicted in laboratory toxicity tests using mean PEC-quotients (USEPA 2000b) 186

List of Figures

Figure 2.1 Recommended procedure for assessing sediment chemistry data 187

Figure 3.1 Relationship between mean PEC quotients and the incidence of toxicity in freshwater toxicity tests (USEPA 2000b) 188

Figure 3.2 Recommended procedure for assessing sediment toxicity data 189

Figure 4.1 Recommended procedure for assessing benthic invertebrate or fish community structure 190

Figure 5.1 Recommended procedure for assessing tissue chemistry data 191

Figure 6.1 Recommended procedure for evaluating fish health data 192

Figure 7.1 The relationship between the mean PEC quotient and the response of *Hyalella azteca* in the 10-day tests (as percent survival) or the response in the Microtox® solid-phase sediment toxicity test (as the EC_{50} expressed as a toxicity reference index). Sediment samples were collected from the Grand Calumet River and Indiana Harbor Canal located in northwestern Indiana (Ingersoll *et al.* 2001b) 193

Figure 7.2 The relationship between the molar concentration of simultaneously extracted metals to acid volatile sulfide (SEM-AVS) and toxic units of metals in the sediment samples. Toxicity of samples was determined using 10-day whole-sediment tests with *Hyalella azteca* (Ingersoll *et al.* 2001b) 194

Figure 7.3. Tri-axial graphs of sediment quality triad data (Canfield *et al.* 1994) 195

Figure A2.1. Metal/aluminum regression lines with the 95% prediction limits (from Carvalho and Schropp 2001) 196

Executive Summary

Traditionally, concerns relative to the management of aquatic resources in freshwater ecosystems have focused primarily on water quality. As such, early aquatic resource management efforts were often directed at assuring the potability of surface water or groundwater sources. Subsequently, the scope of these management initiatives expanded to include protection of instream (i.e., fish and aquatic life), agricultural, industrial, and recreational water uses. While initiatives undertaken in the past twenty years have unquestionably improved water quality conditions, a growing body of evidence indicates that management efforts directed solely at the attainment of surface water quality may not provide an adequate basis for protecting the designated uses of aquatic ecosystems.

In recent years, concerns relative to the health and vitality of aquatic ecosystems have begun to reemerge in North America. One of the principal reasons for this is that many toxic and bioaccumulative chemicals [such as metals, polycyclic aromatic hydrocarbons (PAHs), polychlorinated biphenyls (PCBs), chlorophenols, organochlorine pesticides (OC pesticides), and polybrominated diphenyl ethers]; which are found in only trace amounts in water, can accumulate to elevated levels in sediments. Some of these pollutants, such as OC pesticides and PCBs, were released into the environment long ago. The use of many of these substances has been banned in North America for more than 30 years; nevertheless, these chemicals continue to persist in the environment. Other contaminants enter our waters every day from industrial and municipal discharges, urban and agricultural runoff, and atmospheric deposition from remote sources. Due to their physical and chemical properties, many of these substances tend to accumulate in sediments. In addition to providing sinks for many chemicals, sediments can also serve as potential sources of pollutants to the water column when conditions change in the receiving water system (e.g., during periods of anoxia, after severe storms).

Information from a variety of sources indicates that sediments throughout North America are contaminated by a wide range of toxic and bioaccumulative substances, including metals, PAHs, PCBs, OC pesticides, a variety of semi-volatile organic chemicals (SVOCs), and polychlorinated dibenzo-*p*-dioxins and polychlorinated dibenzofurans (PCDDs and PCDFs). For example, contaminated sediments pose a major risk to the beneficial uses of aquatic ecosystems throughout the Great Lakes basin, including 43 areas of concern (AOCs) that have been identified by the International Joint Commission. The imposition of fish consumption advisories has adversely affected commercial, sport, and food fisheries in many areas. In addition, degradation of the benthic community and other factors have adversely affected fish and wildlife populations. Furthermore, fish in many of these areas have been observed to have higher levels of tumors and other abnormalities than fish from reference areas. Contaminated sediments have also threatened the viability of many commercial ports through the imposition of restrictions on dredging of navigational channels and disposal of dredged materials. Overall, contaminated sediments have been linked to 11 of the 14 beneficial use impairments that have been documented at the Great Lakes AOCs. Such use impairments have also been observed elsewhere in Canada and the United States.

In response to the concerns that have been raised regarding contaminated sediments, responsible authorities throughout North America have launched programs to support the assessment, management, and remediation of contaminated sediments. The information that has been generated under these programs provide important guidance for designing and implementing investigations at sites with contaminated sediments. In addition, guidance has been developed under various sediment-related programs to support the collection and interpretation of sediment quality data. While such guidance has unquestionably advanced the field of sediment quality assessments, the users of the individual guidance documents have expressed a need to consolidate this information into an integrated ecosystem-based framework for assessing and managing sediment quality in freshwater ecosystems (i.e., as specified under the Great Lakes Water Quality Agreement). Practitioners in this field have also indicated the need for additional guidance on the applications of the various tools that support sediment quality assessments. Furthermore, the need for additional guidance on the design of sediment quality monitoring programs and on the interpretation of the resultant data has been identified.

This guidance manual, which comprises a three-volume series, is not intended to supplant the existing guidance on sediment quality assessment. Rather, this guidance manual is intended to further support the design and implementation of assessments of sediment quality conditions by:

- Presenting an ecosystem-based framework for assessing and managing contaminated sediments (Volume I);
- Describing the recommended procedures for designing and implementing sediment quality investigations (Volume II); and,
- Describing the recommended procedures for interpreting the results of sediment quality investigations (Volume III).

The first volume of the guidance manual, ***An Ecosystem-Based Framework for Assessing and Managing Contaminated Sediments in the Freshwater Ecosystems***, describes the five step process that is recommended to support the assessment and management of sediment quality conditions (i.e., relative to sediment-dwelling organisms, aquatic-dependent wildlife, and human health). Importantly, the document provides an overview of the framework for ecosystem-based sediment quality assessment and management (Chapter 2). In addition, the recommended procedures for identifying sediment quality issues and concerns and compiling the existing knowledge base are described (Chapter 3). Furthermore, the recommended procedures for establishing ecosystem goals, ecosystem health objectives, and sediment management objectives are presented (Chapter 4). Finally, methods for selecting ecosystem health indicators, metrics, and targets for assessing contaminated sediments are described (Chapter 5). Together, this guidance is intended to support planning activities related to contaminated sediment assessments, such that the resultant data are likely to support sediment management decisions at the site under investigation. More detailed information on these and other topics related to the assessment and management of contaminated sediments can be found in the publications that are listed in the bibliography (Appendix 2).

The second volume of the series, ***Design and Implementation of Sediment Quality Investigations***, describes the recommended procedures for designing and implementing sediment quality assessment programs. More specifically, an overview of the recommended framework for assessing and managing sediment quality conditions process is presented in this document (Chapter 2). In addition, this volume describes the recommended procedures for conducting preliminary and detailed site investigations to assess sediment quality conditions (Chapters 3 and 4). Furthermore, the factors that need to be considered in the development of sampling and analysis plans for assessing contaminated sediments are described (Chapter 5). Supplemental guidance on the design of sediment sampling programs, on the evaluation of sediment quality data, and on the management of contaminated sediment is provided in the Appendices to this volume. The appendices of this document also describe the types and objectives of sediment quality assessments that are commonly conducted in freshwater ecosystems.

The third volume in the series, ***Interpretation of the Results of Sediment Quality Investigations***, describes the four types of information that are commonly used to assess contaminated sediments, including sediment and pore water chemistry data (Chapter 2), sediment toxicity data (Chapter 3), benthic invertebrate community structure data (Chapter 4), and bioaccumulation data (Chapter 5). Some of the other tools that can be used to support assessments of sediment quality conditions are also briefly described (e.g., fish health assessments; Chapter 6). The information compiled on each of the tools includes: descriptions of its applications, advantages, and limitations; discussions on the availability of standard methods, the evaluation of data quality, methodological uncertainty, and the interpretation of associated data; and, recommendations to guide the use of each of these individual indicators of sediment quality conditions. Furthermore, guidance is provided on the interpretation of data on multiple indicators of sediment quality conditions (Chapter 7). Together, the information provided in the three-volume series is intended to further support the design and implementation of focused sediment quality assessment programs.

List of Acronyms

%	percent
µg	microgram
µg/kg	micrograms per kilogram
µg/L	micrograms per liter
µmol/g	micromoles per gram
AET	apparent effects threshold
AETA	Apparent Effects Threshold Approach
Al	aluminum
ANOVA	analysis of variance
AOC	Area of Concern
APHA	American Public Health Association
ARCS Program	Assessment and Remediation of Contaminated Sediments Program
ASTM	American Society for Testing and Materials
AVS	acid volatile sulfides
BCE	British Columbia Environment
BCWMA	British Columbia Waste Management Act
BEST	biomonitoring of environmental status and trends
BSAF	biota-sediment bioaccumulation factor
CA	Consensus Approach
CAC	citizens advisory committee
CCME	Canadian Council of Ministers of the Environment
CCREM	Canadian Council of Resource and Environment Ministers
CDF	confined disposal facility
CEPA	Canadian Environmental Protection Act
CERCLA	Comprehensive Environmental Response, Compensation, and Liability Act
CERCLIS	Comprehensive Environmental Response, Compensation, and Liability Information System
CI	confidence interval
CLP	Contract Laboratory Program
COC	contaminant of concern
COPC	chemical of potential concern
CRLD	contract required detection limit
CSO	combined sewer overflow
CSR	Contaminated Sites Regulation
CWA	Clean Water Act
-d	- days
DDT	dichlorodiphenyl-trichloroethane
DDTs	*p,p'*-DDT, *o,p'*-DDT, *p,p'*-DDE, *o,p'*-DDE, *p,p'*-DDD, *o,p'*-DDD, and any metabolite or degradation product
DELT	deformities, fin erosion, lesions, and tumors
DL	detection limit

DM	dredged material
DO	dissolved oxygen
DOE	Department of the Environment
DOI	Department of the Interior
DQO	data quality objective
DSI	detailed site investigation
DW	dry weight
EC	Environment Canada
EC_{50}	median effective concentration affecting 50 percent of the test organisms
EEC	European Economic Community
ELA	Effects Level Approach
EMAP	Environmental Monitoring and Assessment Program
EPT	Ephemeroptera, Plecoptera, Trichoptera (i.e., mayflies, stoneflies, caddisflies)
EqPA	Equilibrium Partitioning Approach
ERL	effects range low
ERM	effects range median
EROD	ethoxyresorufin-*O*-deethylase
ESG	equilibrium-partitioning sediment quality guidelines
FCV	final chronic values
FD	factual determinations
FIFRA	Federal Insecticide, Rodenticide and Fungicide Act
gamma-BHC	gamma-hexachlorocyclohexane (lindane)
GFAA	graphite furnace atomic absorption
GIS	geographic information system
-h	- hours
H_2S	hydrogen sulfide
HC	Health Canada
Hcl	hydrochloric acid
IBI	Index of biotic integrity
IC_{50}	median inhibition concentration affecting 50 percent of test organisms
ICP	inductively coupled plasma-atomic emission spectrometry
ID	insufficient data
IDEM	Indiana Department of Environmental Management
IJC	International Joint Commission
IWB	index of well-being
K_{oc}	organic carbon partition coefficients
K_{ow}	octanol-water partition coefficients
Kp	sediment/water partition coefficients
LC_{50}	median lethal concentration affecting 50 percent of the test organism
LCS/LCSDs	laboratory control sample/laboratory control sample duplicates
Li	lithium
LMP	lakewide management plan
LOD	limit of detection
LOEC	lowest observed effect concentration
LRMA	Logistic Regression Modeling Approach

mean PEC-Q	mean probable effect concentration quotient
MESL	MacDonald Environmental Sciences Ltd.
MET	minimal effect threshold
mg/kg	milligrams per kilogram
mg/L	milligrams per liter
mgs	milligrams
mIBI	macroinvertebrate index of biotic integrity
-min	- minutes
mm	millimeter
mm	millimeters
MPRSA	Marine Protection, Research, and Sanctuaries Act
MS/MSDs	matrix spike/matrix spike duplicates
MSD	minimum significant difference
n	number of samples
NAWQA	National Water Quality Assessment
NEPA	National Environmental Policy Act
NG	no guideline available
NH_3	unionized ammonia
NH_4^+	ionized ammonia
NOAA	National Oceanic and Atmospheric Administration
NOEC	no observed effect concentration
NPDES	National Pollutant Discharge and Elimination System
NPL	National Priorities List
NPO	nonpolar organics
NR	not reported
NRDA	natural resource damage assessment
NSQS	National Sediment Quality Survey
NSTP	National Status and Trends Program
NT	not toxic
NYSDEC	New York State Department of Environmental Conservation
OC	organic carbon
OC pesticides	organochlorine pesticides
OECD	Organization of Economic Cooperation and Development
OEPA	Ohio Environmental Protection Agency
OERR	Office of Emergency and Remedial Response
OPA	Oil Pollution Act
OPTTS	Office of Prevention, Pesticides, and Toxic Substances
OSW	Office of Solid Waste
OW	The Office of Water
PAET	probable apparent effects threshold
PAHs	polycyclic aromatic hydrocarbons
PARCC	precision, accuracy, representativeness, completeness, and comparability
PCBs	polychlorinated biphenyls
PCDDs	polychlorinated dibenzo-*p*-dioxins
PCDFs	polychlorinated dibenzofurans
PCS	permit compliance system

PEC	probable effect concentration (consensus-based)
PEC-Q	probable effect concentration quotient
PEL	probable effect level
PEL-HA28	probable effect level for *Hyalella azteca*; 28-day test
PQL	protection quantification limit
PSDDA	Puget Sound Dredged Disposal Analysis
PSEP	Puget Sound Estuary Program
PSI	preliminary site investigation
QA/QC	quality assurance/quality control
QAPP	quality assurance project plan
QHEI	qualitative habitat evaluation index
RAP	remedial action plan
RCRA	Resource Conservation and Recovery Act
REF	reference sediment
RPD	relative percent difference
RRH	rapidly rendered harmless
RSD	relative standard deviation
SAB	Science Advisory Board
SAG	Science Advisory Group
SAP	sampling and analysis plan
SEC	sediment effect concentration
SEL	severe effect level
SEM	simultaneously extracted metals
SEM - AVS	simultaneously extracted metal minus acid volatile sulfides
SETAC	Society of Environmental Toxicology and Chemistry
SLCA	Screening Level Concentration Approach
SMS	sediment management standards
SOD	sediment oxygen demand
SPMD	semipermeable membrane device
SQAL	sediment quality advisory levels
SQC	sediment quality criteria
SQG	sediment quality guideline
SQRO	sediment quality remediation objectives
SQS	sediment quality standard
SSLC	species screening level concentration
SSZ	sediment sampling zone
STP	sewage treatment plant
SVOC	semivolatile organic compound
SVOC	semi-volatile organic chemical
T	toxic
TEC	threshold effect concentration
TEL	threshold effect level
TEL-HA28	threshold effect level for *Hyalella azteca*; 28 day test
TET	toxic effect threshold
TIE	toxicity identification evaluation
TMDL	total maximum daily load

TOC	total organic carbon
tPAH	total polycyclic aromatic hydrocarbons
TRA	Tissue Residue Approach
TRG	tissue residue guideline
TRV	toxicity reference values
TSCA	Toxic Substances Control Act
USACE	United States Army Corps of Engineers
USDOI	United States Department of the Interior
USEPA	United States Environmental Protection Agency
USFWS	United States Fish and Wildlife Service
USGS	United States Geological Survey
VOC	volatile organic compound
WDOE	Washington Department of Ecology
WMA	Waste Management Act
WQC	water quality criteria
WQS	water quality standards
WW	wet weight

Glossary of Terms

Acute toxicity – The response of an organism to short-term exposure to a chemical substance. Lethality is the response that is most commonly measured in acute toxicity tests.

Acute toxicity threshold – The concentration of a substance above which adverse effects are likely to be observed in short-term toxicity tests.

Altered benthic invertebrate community – An assemblage of benthic invertebrates that has characteristics (i.e., mIBI score, abundance of EPT taxa) that are outside the normal range that has been observed at uncontaminated reference sites.

Aquatic ecosystem – All the living and nonliving material interacting within an aquatic system (e.g., pond, lake, river, ocean).

Aquatic invertebrates – Animals without backbones that utilize habitats in freshwater, estuaries, or marine systems.

Aquatic organisms – The species that utilize habitats within aquatic ecosystems (e.g., aquatic plants, invertebrates, fish, amphibians and reptiles).

Benthic invertebrate community – The assemblage of various species of sediment-dwelling organisms that are found within an aquatic ecosystem.

Bioaccumulation – The net accumulation of a substance by an organism as a result of uptake from all environmental sources.

Bioaccumulation-based sediment quality guidelines (SQGs) – Sediment quality guidelines that are established to protect fish, aquatic-dependent wildlife, and human health against effects that are associated with the bioaccumulation of contaminants in sediment-dwelling organisms and subsequent food web transfer.

Bioaccumulative substances – The chemicals that tend to accumulate in the tissues of aquatic and terrestrial organisms.

Bioavailability – Degree to which a chemical can be absorbed by and/or interact with an organism.

Bioconcentration – The accumulation of a chemical in the tissues of an organism as a result of direct exposure to the surrounding medium (e.g., water; i.e., it does not include food web transfer).

Biomagnification – The accumulation of a chemical in the tissues of an organism as a result of food web transfer.

Chemical benchmark – Guidelines for water or sediment quality which define the concentration of contaminants that are associated with low or high probabilities of observing harmful biological effects, depending on the narrative intent.

Chemical of potential concern – A substance that has the potential to adversely affect surface water or biological resources.

Chronic toxicity – The response of an organism to long-term exposure to a chemical substance. Among others, the responses that are often measured in chronic toxicity tests include lethality, decreased growth, and impaired reproduction.

Chronic toxicity threshold – The concentration of a substance above which adverse effects are likely to be observed in long-term toxicity tests.

Congener – A member of a group of chemicals with similar chemical structures (e.g., PCDDs generally refers to a group of 75 congeners that consist of two benzene rings connected to each other by two oxygen bridges).

Consensus-based probable effect concentrations (PECs) – The PECs that were developed from published sediment quality guidelines and identify contaminant concentrations above which adverse biological effects are likely to occur.

Consensus-based threshold effect concentrations (TECs) – The TECs that were developed from published sediment quality guidelines and identify contaminant concentrations below which adverse biological effects are unlikely to occur.

Contaminants of concern (COC) – The substances that occur in environmental media at levels that pose a risk to ecological receptors or human health.

Contaminated sediment – Sediment that contains chemical substances at concentrations that could potentially harm sediment-dwelling organisms, wildlife, or human health.

Conventional variables – A number of variables that are commonly measured in water and/or sediment quality assessments, including water hardness, conductivity, total organic carbon (TOC), sediment oxygen demand (SOD), unionized ammonia (NH_3), temperature, dissolved oxygen (DO), pH, alkalinity

Core sampler – A device that is used to collect both surficial and sub-surface sediment samples by driving a hollow corer into the sediments.

Degradation – A breakdown of a molecule into smaller molecules or atoms.

DELT abnormalities – A number of variables that are measured to assess fish health, including deformities, fin erosion, lesions, and tumors.

Diagenesis – The sum of the physical and chemical changes that take place in sediments after its initial deposition (before they become consolidated into rocks, excluding all metamorphic changes).

Discharge – discharge of oil as defined in Section 311(a)(2) o f the Clean Water Act, and includes, but is not limited to, any spilling, leaking, pumping, pouring, emitting, emptying, or dumping of oil.

Ecosystem – All the living (e.g., plants, animals, and humans) and nonliving (rocks, sediments, soil, water, and air) material interacting within a specified location in time and space.

Ecosystem-based management – An approach that integrates the management of natural landscapes, ecological processes, physical and biological components, and human activities to maintain or enhance the integrity of an ecosystem. This approach places equal emphasis on concerns related to the environment, the economy, and the community (also called the ecosystem approach).

Ecosystem goals – Are broad management goals which describe the long-term vision that has been established for the ecosystem.

Ecosystem metrics – Identify quantifiable attributes of the indicators and defines acceptable ranges, or targets, for these variables.

Ecosystem objectives – Are developed for the various components of the ecosystem to clarify the scope and intent of the ecosystem goals. These objectives should include target schedules for being achieved.

Endpoint – A measured response of a receptor to a stressor. An endpoint can be measured in a toxicity test or in a field survey.

Epibenthic organisms – The organisms that live on the surface of sediments.

Exposure – Co-occurrence of or contact between a stressor (e.g., chemical substance) and an ecological component (e.g., aquatic organism).

Grab (Dredge) samplers – A device that is used to collect surficial sediments through a scooping mechanism (e.g. petite ponar dredge).

Hazardous substance – hazardous substance as defined in Section 101(14) of CERCLA.

Index of biotic integrity (IBI) – A parameter that is used to evaluate the status of fish communities. The IBI integrates information on species composition (i.e., total number of species, types of species, percent sensitive species, and percent tolerant species), on trophic composition (i.e., percent omnivores, percent insectivores, and percent pioneer species), and on fish condition.

Infaunal organisms – The organisms that live in sediments.

Injury – a measurable adverse change, either long or short-term, in the chemical or physical quality or the viability of a natural resource resulting either directly or indirectly from exposure to a discharge of oil or release of a hazardous substance, or exposure to a product of reactions resulting from the discharge to oil or release of a hazardous substance. As used in this part, injury encompasses the phrases "injury", "destruction", and "loss". Injury definitions applicable to specific resources are provided in Section 11.62 of this part (this definition is from the Department of the Interior Natural Resource Damage Assessment Regulations).

Macroinvertebrate index of biotic integrity (mIBI) – The mIBI was used to provide information on the overall structure of benthic invertebrate communities. The scoring criteria for this metric includes such variables as number of taxa, percent dominant taxa, relative abundance of EPT taxa, and abundance of chironomids.

Mean probable effect concentration-quotient (PEC-Q) – A measure of the overall level of chemical contamination in a sediment, which is calculated by averaging the individual quotients for select chemicals of interest..

Natural resources – land, fish, wildlife, biota, air, water, ground water, drinking water supplies, and other such resources belonging to, managed by, held in trust by, appertaining to, or otherwise controlled by the federal government (including the resources of the fishery conservation zone established by the Magnuson Fishery Conservation and Management Act of 1976), State or local government, or any foreign government and Indian tribe. These natural resource have been categorized into the following five groups: surface water resources, ground water resources, air resources, geologic resources, and biological resources.

Natural resources damage assessment – the process of collecting, compiling, and analyzing information, statistics, or data through prescribed methodologies to determine damages for injuries to natural resources as set forth in this part.

Neoplastic – Refers to abnormal new growth.

Oil – oil as defined in Section 311(a)(1) of the Clean Water Act, of any kind or in any form, including, but not limited to, petroleum, fuel oil, sludge, oil refuse, and oil mixed with wastes other that dredged spoil.

Piscivorus wildlife species – The wildlife species that consume fish as part of all of their diets (e.g., herons, kingfishers, otter, osprey, and mink).

Population – An aggregate of individual of a species within a specified location in time and space.

Pore water – The water that occupies the spaces between sediment particles.

Probable effect concentration (PEC) – Concentration of a chemical in sediment above which adverse biological effects are likely to occur.

Probable effect concentration-quotient (PEC-Q) – A PEC-Q is a measure of the level of chemical contamination in sediment relative to a sediment quality guideline, and is calculated by dividing the measured concentration of a substance in a sediment sample by the corresponding PEC.

Receptor – A plant or animal that may be exposed to a stressor.

Release – A release of a hazardous substance as defined in Section 101(22) of CERCLA.

Sediment – Particulate material that usually lies below water.

Sediment-associated contaminants – Contaminants that are present in sediments, including whole sediments or pore water.

Sediment chemistry data – Information on the concentrations of chemical substances in whole sediments or pore water.

Sediment-dwelling organisms – The organisms that live in, on, or near bottom sediments, including both epibenthic and infaunal species.

Sediment injury – The presence of conditions that have injured or are sufficient to injure sediment-dwelling organisms, wildlife, or human health.

Sediment quality guideline – Chemical benchmark that is intended to define the concentration of sediment-associated contaminants that is associated with a high or a low probability of observing harmful biological effects or unacceptable levels of bioaccumulation, depending on its purpose and narrative intent.

Sediment quality targets – Chemical or biological benchmarks for assessing the status of each metric.

Simultaneously extracted metals (SEM) – Divalent metals - commonly cadmium, copper, lead, mercury, nickel, and zinc - that form less soluble sulfides than does iron or manganese and are solubilized during the acidification step (0.5m HCl for 1 hour) used in the determination of acid volatile sulfides in sediments.

Stressor – Physical, chemical, or biological entities that can induce adverse effects on ecological receptors or human health.

Surface water resources – The waters of North America, including the sediments suspended in water or lying on the bank, bed, or shoreline and sediments in or transported through coastal and marine areas. This term does not include ground water or water or sediments in ponds, lakes, or reservoirs designed for waste treatment under the Resource Conservation and Recovery Act of 1976 (RCRA), 42 U.S.C. 6901-6987 or the Clean Water Act, and applicable regulations.

Threshold effect concentration (TEC) – Concentration of a chemical in sediment below which adverse biological effects are unlikely to occur.

Tissue – A group of cells, along with the associated intercellular substances, which perform the same function within a multicellular organism.

Tissue residue guideline (TRG) – Chemical benchmark that is intended to define the concentration of a substance in the tissues of fish or invertebrates that will protect fish-eating wildlife against effects that are associated with dietary exposure to hazardous substances.

Trophic level – A portion of the food web at which groups of animals have similar feeding strategies.

Trustee – Any Federal natural resources management agency designated in the National Contingency Plan and any State agency designated by the Governor of each State, pursuant to Section 107(f)(2)(B) of CERCLA, that may prosecute claims for damages under Section 107(f) or 111(b) of CERCLA; or any Indian tribe, that may commence an action under Section 126(d) of CERCLA.

Wildlife – The fish, reptiles, amphibians, birds, and mammals that are associated with aquatic ecosystems.

Whole sediment – Sediment and associated pore water.

Acknowledgments

The authors would like to acknowledge the efforts of a number of individuals who contributed to the preparation of '*A Guidance Manual to Support the Assessment of Contaminated Sediments in Freshwater Ecosystems*'. First, we would like to thank the members of the Science Advisory Group on Sediment Quality Assessment for their insight and guidance on the need for and elements of this Guidance Manual. We would also like to thank the instructors of the various short courses on sediment quality assessment for providing access to instructional materials that provided a conceptual basis for many of the sections included in the Guidance Manual. Furthermore, we would like to express our sincerest appreciation to the members of the project Steering Committee for providing oversight and excellent review comments on previous drafts of this report. The Steering Committee consisted of the following individuals:

Tom Balduf (Ohio Environmental Protection Agency)
Peter Landrum (National Oceanic and Atmospheric Administration)
Judy Crane (Minnesota Pollution Control Agency)
Scott Cieniawski (United States Environmental Protection Agency)
Karen Woodfield (New York State Department of Environmental Conservation)
Bonnie Eleder (United States Environmental Protection Agency)
Roger Jones (Michigan Department of Environmental Quality)
John Estenik (Ohio Environmental Protection Agency)
Walter Berry (United States Environmental Protection Agency)
Scott Ireland (United States Environmental Protection Agency)
Demaree Collier (United States Environmental Protection Agency)
Marc Tuchman (United States Environmental Protection Agency)
Frank Estabrooks (New York State Department of Environmental Conservation)
Jay Field (National Oceanic and Atmospheric Administration)
Lee Liebenstein (Wisconsin Department of Natural Resources)
Jan Miller (United States Army Corps of Engineers)
T.J. Miller (United States Fish and Wildlife Service)
Dave Mount (United States Environmental Protection Agency)
Gail Sloane (Florida Department of Environmental Protection Agency)
Eric Stern (United States Environmental Protection Agency)
William Creal (Michigan Department of Environmental Quality)
Kelly Burch (Pennsylvania Department of Environmental Protection)
Mike Macfarlane (British Columbia Ministry of Water, Air and Land Protection)

Finally, timely review comments on the final draft of the Guidance Manual were provided by Scott Cieniawski, Demaree Collier, Marc Tuchman, Jay Field, Judy Crane, and Mike Macfarlane. Development of this Guidance Manual was supported in part by the United States Environmental Protection Agency's Great Lakes National Program Office through the grant "Development of a Guidance Manual for Sediment Assessment", Grant Number GL995632-01, awarded to the Sustainable Fisheries Foundation. Additional funding to support the preparation of this report was provided by the Florida Department of Environmental Protection and the British Columbia Ministry of Water, Air and Land Protection. This report has been reviewed in accordance with United States Environmental Protection Agency, United States Geological Survey, and Sustainable Fisheries Foundation policies.

Chapter 1. Introduction

1.0 Background

In response to the concerns that have been raised regarding contaminated sediments, a number of programs have been established or expanded to support the assessment and management of contaminated sediments in the United States and Canada (Appendix 1 of Volume III). The information generated under these programs provides important guidance for designing and implementing investigations at sites with contaminated sediments (see USEPA 1994; MacDonald 1994a; 1994b; Reynoldson *et al.* 2000; Ingersoll *et al.* 1997; USEPA and USACE 1998a; ASTM 2001a; USEPA 2000a; Krantzberg *et al.* 2001). While these guidance documents have unquestionably advanced the field of sediment quality assessment, the users of these individual guidance documents have expressed a need to consolidate this information into an integrated ecosystem-based framework for assessing and managing sediment quality in freshwater ecosystems.

This guidance manual, which comprises a three-volume series, is not intended to supplant the existing guidance documents on sediment quality assessment (e.g., USEPA 1994; Reynoldson *et al.* 2000; USEPA and USACE 1998a; USEPA 2000a; ASTM 2001a; Krantzberg *et al.* 2001). Rather, this guidance manual is intended to further support the design and implementation of assessments of sediment quality conditions by:

- Presenting an ecosystem-based framework for assessing and managing contaminated sediments (Volume I);
- Describing the recommended procedures for designing and implementing sediment quality investigations (Volume II); and,
- Describing the recommended procedures for interpreting the results of sediment quality investigations (Volume III).

The first volume of the guidance manual, ***An Ecosystem-Based Framework for Assessing and Managing Contaminated Sediments in Freshwater Ecosystems***, describes the five step process that is recommended to support the assessment and management of sediment quality conditions (i.e., relative to sediment-dwelling organisms, aquatic-dependent wildlife, and human health). Importantly, the document provides an overview of the framework for ecosystem-based sediment quality assessment and management (Chapter 2). The recommended procedures for identifying sediment quality issues and concerns and compiling the existing knowledge base are described (Chapter 3). Furthermore, the recommended procedures for establishing ecosystem goals, ecosystem health objectives, and sediment management objectives are presented (Chapter 4). Finally, methods for selecting ecosystem health indicators, metrics, and targets for assessing contaminated sediments are described (Chapter 5). Together, this guidance is intended to support planning activities related to contaminated sediment assessments, such that the resultant data are likely to support sediment management decisions at the site under investigation.

The second volume of the series, ***Design and Implementation of Sediment Quality Investigations***, describes the recommended procedures for designing and implementing sediment quality assessment programs. More specifically, an overview of the recommended framework for assessing and managing sediment quality conditions process is presented in this document (Chapter 2). In addition, this volume describes the recommended procedures for conducting preliminary and detailed site investigations to assess sediment quality conditions (Chapters 3 and 4). Furthermore, the factors that need to be considered in the development of sampling and analysis plans for assessing contaminated sediments are described (Chapter 5). Supplemental guidance on the design of sediment sampling programs, on the evaluation of sediment quality data, and on the management of contaminated sediment is provided in the Appendices to this volume. The appendices of this document also describe the types and objectives of sediment quality assessments that are commonly conducted in freshwater ecosystems.

The third volume in the series, ***Interpretation of the Results of Sediment Quality Investigations***, describes the four types of indicators that are commonly used to assess contaminated sediments, including sediment and pore water chemistry data (Chapter 2), sediment toxicity data (Chapter 3), benthic invertebrate community structure data (Chapter 4), and bioaccumulation data (Chapter 5). Some of the other indicators that can be used to support assessments of sediment quality conditions are also described (e.g., fish health assessments; Chapter 6). The information compiled on each of the indicators includes: descriptions of its applications, advantages, and limitations; discussions on the availability of standard methods, the evaluation of data quality, methodological uncertainty, and the interpretation of associated data; and, recommendations to guide its use. Furthermore, guidance is provided on the interpretation of data on multiple indicators of sediment quality conditions (Chapter 7). Together, the information provided in the three-volume series is intended to further support the design and implementation of focused sediment quality assessment programs.

Chapter 2. Assessment of Sediment and Pore Water Chemistry

2.0 Introduction

Sediment chemistry data represent a fundamental element of sediment quality assessments that are focused on evaluation of the effects of toxic and bioaccumulative substances. Therefore, sediment chemistry is routinely selected as one of the key ecosystem health indicators in most sediment quality investigations (see Volume I for information on the selection of the ecosystem health indicators). To be effective, however, metrics and associated targets must be selected that are relevant to the site under investigation (i.e., relative to the management objectives established; see Chapters 4 and 5 of Volume I). In general, the metrics that are selected for evaluating sediment chemistry typically include the concentrations of the chemicals of potential concern (COPCs) that have been identified for the site. Sediment quality targets are usually identified by selecting sediment quality guidelines (SQGs) that apply to the receptors of concern and desired level of protection at the site. This chapter is intended to provide guidance on the selection of metrics and target for sediment chemistry that will provide the information needed to effectively assess sediment quality conditions at contaminated sites. A description of the recommended uses of SQGs is provided in Appendix 1 of Volume III.

2.1 Selection of Metrics and Targets for Sediment Chemistry

Several types of information can be used to support the selection of appropriate metrics for sediment chemistry. First, current and historic land and water use activities in the vicinity of the site should be determined (see Volume II for more information). Historical data should include information on the nature and location of industrial developments (and associated management practices that could lead to releases of chemical substances) and municipal infrastructure (combined sewer overflows, sewage treatment plants), on the nature and location of any spills that have occurred, and on the nature and general location of non-point pollution sources. In addition, information on the location, composition, and volumes of storm water and effluent discharges is useful for identifying the chemicals that have been or may have been released into surface waters near the site. Evaluation of the environmental fate of these chemicals provides a basis for identifying the substances that are likely to partition into sediments. Finally, existing sediment chemistry data should be assembled and used to identify the chemicals that have been measured at elevated levels (i.e., compared to SQGs or targets) in surficial (i.e., top 10 cm) and deeper sediments. Together, this information can be used to develop a list of COPCs at the site. This list of COPCs can then be used to establish the primary metrics for sediment chemistry at the site. Additional metrics, such as total organic carbon (TOC), grain size, acid volatile sulfides (AVS), ammonia, and hydrogen sulfide should also be included to support interpretation of the resultant data for the primary metrics. The final list of chemical analytes to be measured is also influenced by the equipment, technology, facilities, and funds that are available for the project (see Chapter 3 of Volume I for more information on the identification of COPCs).

The chemicals that are typically analyzed in whole-sediment samples collected near urbanized and industrial areas include trace metals, polycyclic aromatic hydrocarbons (PAHs), polychlorinated biphenyls (PCBs), and several other organic constituents [e.g., polychlorinated dibenzo-*p*-dioxins and polychlorinated dibenzofurans (PCDDs/PCDFs); chlorophenols, and phthalates]. In areas that may be affected by inputs from agricultural activities, it may be appropriate to measure the concentrations of pesticides [such as organochlorines (OCs), carbamates, and organophosphates] in sediment samples. Chemical concentrations are generally reported on a dry weight basis, based on the results of total extraction of sediment samples. However, several other measures of sediment chemistry have also been utilized in various assessments. For example, the concentrations of non-ionic organic contaminants may be normalized to TOC concentrations in sediment (Swartz *et al.* 1987; Di Toro *et al.* 1991). In addition, AVS-normalization procedures may be used to interpret data on the levels of simultaneously extracted metals (SEMs; Di Toro *et al.* 1992; Ankley *et al.* 1996). Furthermore, chemical concentrations can be normalized to percent fines. These normalization procedures are intended to better define the bioavailable fraction of the substance under consideration.

Pore water is the water that occupies the spaces between sediment particles. Pore water can be isolated from the sediment matrix to conduct toxicity testing or to measure the concentrations of chemical substances. ASTM (2001a) and USEPA (2000a) describe procedures for isolating pore water from whole-sediment samples. Evaluation of the concentrations of contaminants in pore water is important because sediment-dwelling organisms are directly exposed to the substances that occur in this sediment phase. For this reason, pore water assessments can provide useful information on the potential effects of sediment-associated contaminants, particularly on infaunal species (i.e., those species that utilize habitats within the sediment matrix). Importantly, the toxicity of sediments to aquatic organisms has been correlated to the concentrations of contaminants in pore water (Di Toro *et al.* 1991; Ankley *et al.* 1996). Contaminants in pore water also represent hazards to water column species because these contaminants can be transported into overlying waters through chemical partitioning, diffusion, bioturbation, or resuspension processes. However, data on the concentrations of chemicals in pore water may not fully represent the total exposure of sediment-dwelling organisms to sediment-associated contaminants, particularly for compounds with higher octanol-water partition coefficients (K_{ow}s) that bind strongly to organic carbon in the sediment (Harkey *et al.* 1994). For this reason, pore water chemistry alone should not be used to evaluate total exposure to sediment-associated contaminants.

Selection of appropriate metrics for pore water chemistry should be done in a manner that is consistent with the process used to select the metrics for whole-sediment chemistry. In addition to the substances that are expected to partition into sediments (due to their physical-chemical properties), it may be appropriate to include additional COPCs that are likely to partition primarily into water. It is necessary to include a number of variables (e.g., pH, water temperature, water hardness, dissolved oxygen) that will provide ancillary information for interpreting the data on the primary chemical metrics.

Sediment chemistry data provide information that is directly relevant for determining if sediments within an assessment area are contaminated with toxic and/or bioaccumulative substances. However, information on the concentrations of contaminants in whole sediments (i.e., the metrics for sediment chemistry) does not, by itself, provide a basis for determining if the ecosystem goals and objectives are being achieved. For this reason, establishing sediment quality targets that define the levels of each metric is necessary (i.e., the COPCs and mixtures of COPCs) that are likely to support the designated uses of the aquatic ecosystem (i.e., the benthic invertebrate community). These targets can be established by selecting appropriate SQGs for each COPC at the site. Such SQGs can be derived using information on contemporary background levels and/or on the concentrations associated with a pre-selected probability of observing adverse biological effects (Field *et al.* 2002; Appendices 2 and 3 of Volume III).

Effects-based SQGs represent the principal tools for establishing sediment quality targets that correspond to the specific management goals that have been established for the site under consideration. A variety of numerical SQGs have been developed to support sediment quality assessments in North America (Tables 2.1 and 2.2; Appendix 3 of Volume III). The approaches selected by individual jurisdictions depend on the receptors that are to be considered (e.g., sediment-dwelling organisms, wildlife, or humans), the degree of protection that is to be afforded, the geographic area to which the values are intended to apply (e.g., site-specific, regional, or national), and their intended uses (e.g., screening tools, remediation objectives, identifying toxic and not toxic samples, bioaccumulation assessment).

Guidelines for assessing sediment quality relative to the potential for adverse effects on sediment-dwelling organisms in freshwater systems have been derived using a combination of theoretical and empirical approaches, primarily including the equilibrium partitioning approach [(EqPA) which is used to develop equilibrium-partitioning SQGs (ESGs); Di Toro *et al.* 1991; NYSDEC 1999; USEPA 1997], screening level concentration approach (SLCA; Persaud *et al.* 1993), effects range approach (ERA; Long and Morgan 1991; USEPA 1996), effects level approach (ELA; Smith *et al.* 1996; USEPA 1996), the apparent effects threshold approach (AETA; Cubbage *et al.* 1997), and the consensus-based approach (Swartz 1999; MacDonald *et al.* 2000a; 2000b; USEPA 2000b; Ingersoll *et al.* 2001a). Application of these methods has resulted in the derivation of numerical SQGs for many COPCs in freshwater sediments (Tables 2.1 and 2.2; Appendix 3 of Volume III).

In addition to causing direct effects on aquatic biota (Chapters 3 and 4 of Volume III), sediment-associated contaminants can accumulate in the tissues of sediment-dwelling organisms (Chapter 5 of Volume III). Because many benthic and epibenthic species represent important components of the food web, such contaminants can be transferred to higher trophic levels in the food web. In this way, contaminated sediments represent a potential hazard to the wildlife species that consume aquatic organisms. As such, sediment chemistry represents an important ecosystem health indicator with respect to the potential for effects on aquatic-dependent wildlife species.

The concentrations of bioaccumulative substances in sediments represent the primary metrics for assessing sediment chemistry relative to aquatic-dependent wildlife (Chapter 5 of Volume III). In general, the target analytes in whole sediments should be selected based on historic information on water and land uses in the vicinity of the site under investigation, as well as a review of existing sediment and tissue chemistry data. The bioaccumulative substances that are commonly measured in whole-sediment samples collected in the vicinity of urban, industrial, and agricultural areas include certain PAHs, PCBs, OC pesticides, chlorophenols, certain trace metals (e.g., mercury), and PCDDs/PCDFs (ASTM 2001a; USEPA 2000a).

Residue-based SQGs provide practical tools for establishing targets for sediment chemistry relative to the potential for bioaccumulation (Cook *et al.* 1992; Appendix 3 of Volume III). Residue-based SQGs define the maximum concentrations of individual chemicals or classes of chemicals in sediments that are predicted to result in tolerable levels of those substances in the tissues of aquatic organisms (i.e., below the levels associated with adverse effects in piscivorus wildlife). The first step in the development of residue-based SQGs involves the derivation or selection of an appropriate tissue residue guideline (TRG) for the substance or substances under consideration (e.g., the New York State Department of Environmental Conservation fish flesh criteria for piscivorus wildlife - Newell *et al.* 1987). Subsequently, relationships between concentrations of contaminants in sediments and contaminant residues in aquatic biota needs to be established. In general, the necessary biota-sediment accumulation factors (BSAFs) are determined from field studies, based on the results of bioaccumulation tests, and/or estimated using various modeling approaches. The SQGs are then derived by dividing the TRG by the BSAF (Cook *et al.* 1992; NYSDEC 1999). Because it is difficult to accurately predict relationships between sediment chemistry and the concentrations of COPCs in the tissues of aquatic organisms, potential risks of piscivorus wildlife identified using the SQGs should be confirmed using site-specific tissue residue data and appropriate TRGs.

Contaminated sediment represents a significant environmental concern with respect to the protection of human health. Humans can be directly exposed to contaminated sediments through primary contact recreation, including swimming and wading in affected waterbodies. In addition, indirect exposures to sediment-associated contaminants can occur when human consume fish, shellfish, or wildlife tissues that have become contaminated due to bioaccumulation in the food web (Crane 1996). Therefore, sediment chemistry represents an important ecosystem health indicator for assessing the potential effects of COPCs on human health. The bioaccumulation-based SQGs for the protection of human health that were developed by New York State Department of Environmental Conservation (NYSDEC 1999) and Washington State Department of Health (1995; 1996) provide a basis for establishing sediment quality targets relative to the protection of human health.

2.2 Availability of Standard Methods

Standard methods have been developed to support the characterization of whole-sediment or pore water samples for most major COPCs (i.e., by American Society for Testing and Materials, United States Environmental Protection Agency, Organization for Economic Cooperation and Development, Environment Canada; Appendix 4 of Volume III). In addition, methods used to develop and evaluate SQGs have been described in the peer-reviewed literature (Appendix 3 of Volume III).

2.3 Advantages and Disadvantages of Sediment Chemistry Data

One of the principal strengths of using sediment chemistry data for whole sediments in assessing the potential effects on sediment-dwelling organisms is that it provides direct information on the presence and concentrations of COPCs in sediments (Table 2.3). In addition, standard methods have been established for determining the concentrations of many analytes in whole-sediment samples. Because measurements of sediment chemistry can be both accurate and precise, they provide a reliable basis for discriminating between contaminated and uncontaminated sites. Furthermore, analytical methods have been developed that may provide information on the potential bioavailability of certain substances (e.g., SEM minus AVS and organic carbon normalization of non-ionic organic compounds). Importantly, reliable SQGs have been developed for many COPCs, which provide a basis of interpreting sediment chemistry data relative to the potential for effects on sediment-dwelling organisms.

One of the main limitations of sediment chemistry data is that, by itself, it can not provide a basis for assessing the potential effects of contaminated sediments. The utility of these data may also be limited by the suite of analytes selected for determination. For example, important chemicals may be missed if the available land and water use data are not collected and appropriately interpreted (e.g., PCDDs /PCDFs should be measured in the vicinity of pulpmills, pesticides should be measured near agricultural areas). In some cases, the utility of these data is also limited by the inappropriate use of analytical methods (i.e., which do not support achievement of target detection limits) or by inadequate quality assurance practices (i.e., such that evaluating the reliability of the data is not possible).

One of the strengths of pore water chemistry data is that it provides information on the levels of chemical contaminants in this important exposure medium (Table 2.3). As such, pore water chemistry data facilitates the identification of the substances that are causing or substantially contributing to any adverse biological effects that are observed. As is the case for whole-sediment chemistry, standard methods have been established for determining the concentrations of many COPCs in pore water. Importantly, measurements of the concentrations of COPCs in pore water provide direct information on the sediment-associated contaminant fraction that is likely to be most available to sediment-dwelling organisms.

Pore water chemistry data also have a number of limitations that restrict their application in sediment quality assessments. First, pore water chemistry data cannot be used alone to evaluate the potential for effects on sediment-dwelling organisms (i.e., companion tools are needed to link contaminant concentrations to the effects on various receptors). Second, the procedures that are used to obtain pore water from whole sediments have the potential to alter pore water chemistry. Third, obtaining sufficient volumes of pore water to support analysis of a full suite of chemical analytes is often difficult, particularly when low detection limits are required to assess risks associated with exposures of sediment-dwelling organisms to organic contaminants. Pore water chemistry can also vary temporally (e.g., seasonally). Finally, the utility of these data can be difficult to evaluate due to use of inappropriate methods or inadequate quality assurance practices (ASTM 2001a; USEPA 2000a). Measuring water quality characteristics of the pore water to assist in the interpretation of these data is important (i.e., hardness, alkalinity, pH, dissolved organic carbon).

Interpretation of sediment chemistry data relative to the potential for effects on wildlife species is complicated by differences in BSAFs and food web transfer rates among sites. As such, predictions of contaminant accumulation rates from sediment to biota should generally be validated using appropriate field and/or laboratory procedures. Residue-based SQGs represent important tools for conducting sediment quality assessments for several reasons. First and foremost, residue-based SQGs explicitly consider the potential for bioaccumulation and effects on higher trophic levels. In addition, the residue-based SQGs provide a basis for interpreting sediment chemistry data, in terms of the potential for adverse effects on wildlife. Such assessments should be supported by direct measurements of contaminant concentrations in the tissues of aquatic organisms and wildlife species to assure that actual hazards are identified (Chapter 5 of Volume III).

One of the disadvantages of utilizing sediment quality as an indicator of effects on wildlife is that TRGs for the protection of wildlife have not been developed for many COPCs (Newell *et al.* 1987; Cook *et al.* 1992). Therefore, SQGs for such COPCs must be developed before effects on aquatic-dependent wildlife can be assessed using sediment chemistry data.

When considered in conjunction with food web models, sediment chemistry data can be used to predict the concentrations of contaminants in fish, shellfish, and wildlife tissues; hence, it is possible to evaluate various human health exposure scenarios associated with the consumption of contaminated tissues. The availability of standard analytical methods, procedures for assessing data quality (i.e., accuracy, precision, detection limits), and procedures for evaluating the bioavailability of sediment-associated contaminants make sediment chemistry a reliable indicator of sediment quality conditions.

In spite of the advantages noted above, interpretation of sediment chemistry data relative to the potential for effects on human health poses a challenge for several reasons. First, sediment chemistry data, by themselves, cannot be used to evaluate the potential for effects on human health. Interpretation of such data relative to human health necessarily requires effects-based SQGs. Relative to direct contact recreation, derivation of such guidelines necessitates the development of exposure scenarios that are relevant to the site under investigation (i.e., in addition to appropriate toxicological data). Second, estimation of the levels of bioaccumulative substances in the tissues of fish, shellfish, or wildlife necessitates the use of bioaccumulation models, which may or may not be directly applicable to the ecosystem under study. Furthermore, the actual exposures of humans to contaminated tissues can be reduced through the imposition of fish consumption advisories. Therefore, effects on human health that are predicted based on sediment chemistry data may not actually be observed in the field.

2.4 Evaluation of Data Quality

The use of performance-based methods has been recommended for sediment toxicity testing (ASTM 2001a; USEPA 2000a). Performance-based methods provide investigators with a higher degree of confidence that project data quality objectives (DQOs) will be met. This approach is also highly relevant for guiding the generation of sediment chemistry data. In this context, performance standards should be established for data accuracy, data precision, and analyte detection limits. Guidance on the establishment of DQOs and evaluation of data quality is provided in Appendix 3 of Volume II. Importantly, target detection limits should be established at concentrations lower than the selected sediment quality target (i.e., below a selected SQG). Appendix 4 of Volume III outlines criteria that should be considered when evaluating the quality of chemistry data used in an assessment of sediment quality. A quality assurance project plan (QAPP) should be developed to describe the experimental design and sampling procedures for sediment collection and chemical analyses.

2.5 Methodological Uncertainty

A review of uncertainty associated with endpoints commonly used in sediment ecological risk assessments and approaches for addressing these sources of uncertainty was provided by Ingersoll *et al.* (1997). Endpoints included in this evaluation included: toxicity tests (both the fraction tested and the endpoints selected); benthic invertebrate assessments; bioaccumulation assessments; sediment chemistry; and, sediment chemistry and SQGs. A series of criteria were established by Ingersoll *et al.* (1997) to support consistent assessments of the uncertainty associated with each of these measurement endpoints. These evaluation criteria included: precision; ecological relevance; causality; sensitivity; interferences; standardization; discrimination; bioavailability; and, field validation.

The results of these evaluations are presented in Table 2.4 for sediment chemistry and in Table 2.5 for SQGs. Uncertainty associated with lack of knowledge is indicated with an asterisk in these tables to differentiate it from systematic uncertainty, which can be rectified (methodologically) or quantified (sampling decisions and design).

2.5.1 Uncertainty Associated with Sediment Chemistry

The uncertainty associated with the following measures of sediment chemistry were evaluated by Ingersoll *et al.* (1997; Table 2.4):

- Bulk sediment analysis using total extraction of sediments;
- Normalization of non-ionic organic contaminants to TOC concentration of sediment;
- Metal speciation as derived by AVS or by evaluating other partitioning phases;
- Concentration of contaminants in pore water samples;
- Concentrations of contaminants in elutriate samples; and,

- Concentrations of reference elements (which are regional reference levels to which contaminant concentrations are compared).

The evaluation performed by Ingersoll *et al.* (1997) addresses the uncertainty associated with the use of sediment chemistry alone in sediment assessments. A lower level of uncertainty would be assigned to several of the chemistry measures if these endpoints were used in combination with other endpoints (e.g., toxicity tests, benthic community assessments).

Precision was defined by Ingersoll *et al.* (1997) in terms of the robustness of the analytical method. That is, procedures that generate similar concentrations in repeated analyses of the same samples were considered to have a lower level of uncertainty than those that generate variable results. The lowest level of uncertainty was assigned to bulk sediment, TOC-normalization, SEM-AVS (i.e., on a molar basis), elutriate, and reference element measurements because a high level of precision can be attained using existing analytical methods. Pore water chemistry and procedures intended to determine the species of contaminant present in the sample (speciation procedures) were assigned a higher level of uncertainty resulting from the lack of routine methods used in these analyses. Ecological relevance was evaluated in terms of linkages to receptors that are to be protected. In this respect, bulk sediment chemistry, elutriates, and reference element measurements were rated low since these approaches are not based on measures of bioavailability or are not direct measures of ecological relevance. Total organic carbon normalization, SEM-AVS, metal speciation, and pore water measures were rated as having a moderate level of uncertainty since these measures are based on the principle of evaluating the bioavailable fraction of a chemical in sediment.

Determination of causality (i.e., correctly identifying stressors) was evaluated in terms of the ability of various indicators to determine specific linkages to a COPC, to COPC mixtures, or to sources of chemical contaminants. Low uncertainty was assigned to all of the measures of sediment chemistry, except those which determined chemical concentrations in sediment elutriates. Preparation of elutriates alters the sediment sample, increasing the uncertainty in the sediment contaminant concentration. Although pore water concentrations provide more direct linkages to bulk sediment chemistry, the procedures used to isolate pore water may also introduce considerable uncertainty. Bulk sediment chemistry and reference element-based procedures were considered to provide useful measures for evaluating contaminant sources particularly for certain classes of organics (e.g., PAHs) and for metals. In contrast, elutriate chemistry provides limited information regarding the chemical composition of sediments *in situ* or contaminant sources.

Sensitivity is important because there is a need to reliably identify sediments with high, moderate, and low contaminant concentrations (i.e., as compared to SQGs). Most analytical methods for determining chemical concentrations in sediments are very sensitive. Interferences are considered to be factors which impair accurate determination or interpretation of the concentrations of contaminants in sediment samples. In most cases, interferences are related to sample matrix problems and are analyte specific in any of the categories listed in Table 2.4. Interpretation interferences include particle size variability and anomalously high concentrations of natural sediment components which equilibrate with high concentrations of contaminants.

Standard methods have been developed for virtually all of the analytical procedures outlined in Table 2.4 (e.g., bulk sediment chemistry, pore water chemistry, TOC). However, there are still few methods available which can effectively speciate metals and metalloids in oxidized sediments or can be used to measure non-priority pollutants. Analytical methods are very good discriminators (i.e., establish a gradient) among samples. However, the interpretational uncertainties described above for bulk sediments add substantial uncertainties relative to the discrimination of contamination using this method. Although whole sediment contaminant concentrations do not explicitly intend to quantify the bioavailable fraction, they have been shown to be predictive of biological responses (Ingersoll *et al.* 2001a). The TOC- and AVS-normalization procedures are intended to reduce the level of uncertainty about the bioavailability of non-ionic organics and metals, respectively; however, these procedures have not been shown to increase predictive ability beyond that which has been achieved using whole-sediment chemistry data (Long *et al.* 1998a; Field *et al.* 1999; USEPA 2000b). Elutriate preparation tends to alter bioavailability in unpredictable ways and, therefore, increases uncertainty.

Field validation was interpreted by Ingersoll *et al.* (1997) in terms of the accuracy of the method. That is, the uncertainty about the extent to which measurements of sediment chemistry reflect actual field concentrations of contaminants was evaluated. Bulk sediment chemistry and reference element concentrations have low uncertainty with respect to accuracy because these methods have well-established quality assurance and quality control procedures. A number of uncertainties are associated with the analysis of inorganics (i.e., AVS or metal speciation) and with elutriates (e.g., alterations of the sediments which organisms are exposed to *in situ*, resulting from sample collection, storage, laboratory treatment or other methodological procedures).

2.5.2 Uncertainties Associated with Uses of Sediment Quality Guidelines

Ingersoll *et al.* (1997) placed SQGs into seven categories (Table 2.5):

- Equilibrium-partitioning SQGs (ESGs);
- Effects range low (ERLs) and effects range median (ERMs; threshold and probable effect levels (TELs and PELs) were considered to be functionally similar to the ERLs and ERMs);
- Apparent effects thresholds (AETs);
- Screening level concentrations (SLCs);
- Simultaneously extracted metals minus acid volatile sulfide (SEM-AVS);
- Toxic units models; and,
- Residue-based SQGs (Appendix 3 of Volume III).

Consensus-based SQGs had not been developed or evaluated at the time that the Ingersoll *et al.* (1997) study was conducted (Swartz 1999; MacDonald *et al.* 2000a; 2000b; USEPA 2000a; Ingersoll *et al.* 2001a).

Precision was evaluated by Ingersoll *et al.* (1997) as a measure of the applicability of the SQGs across geographic areas. In terms of precision, the lowest level of uncertainty was assigned to the ESGs because of the extensive toxicology database on which they were derived. Higher uncertainty was assigned to AETs and SLCs because of the site-specificity associated with their derivation. A moderate level of uncertainty was also assigned to the SEM-AVS based guidelines because of the micro-spatial distribution of AVS. Ecological relevance was evaluated in terms of its linkage to the receptors that are to be protected. Guidelines which directly consider mixtures were assigned a relatively low level of uncertainty (mixture models, SEM-AVS guidelines, and the ERL/ERM guidelines derived using data from the field which included contaminant mixtures). Individual ESG values do not consider the effects of mixtures of COPCs and, hence, were assigned a moderate level of uncertainty. Similarly, AETs were assigned a moderate level of uncertainty because of their inherent potential for incorrectly identifying toxic samples as not toxic (i.e., false negatives). The SLCs reflect the lower bound of ecologically relevant sediment concentrations (i.e., background concentrations), but may not necessarily define actual effect concentrations (i.e., false positives; non-toxic samples identified as toxic). Although the TRGs with which the residue-based SQGs were derived are considered to be highly ecologically relevant, more uncertainty is associated with the models which are used to determine the BSAFs.

The ESGs, SEM-AVS, and mixture models were assigned low uncertainty relative to establishing causality because these guidelines are directly derived from experimental determinations of effects of specific chemicals. In contrast, ERLs and ERMs, AETs, and SLCs were assigned higher levels of uncertainty because these guidelines are derived primarily from field observations in which cause and effect relationships were equivocal (i.e., the sediments contained mixtures of contaminants and, hence, determining the identity of the causative agents directly is difficult). Sensitivity was evaluated relative to estimating relatively low contaminant concentrations (i.e., minimize false negatives while allowing for a higher probability of false positives). Optimizing sensitivity (e.g., minimize false negatives) needs to be balanced with ecological relevance (e.g., minimize both false positives and false negatives). Low uncertainty with respect to sensitivity was assigned to the ERLs and SLCs because they tend to be the lowest SQGs. Most of the other SQGs were considered to have a higher level of uncertainty because they are generally higher values (e.g., ESGs, ERMs, and SEM-AVS). The AETs were assigned a high level of uncertainty with respect to sensitivity since they only increase with the addition of new data, making them particularly prone to false negatives. In contrast, the residue-based SQGs were considered to have a lower level of uncertainty because the TRGs upon which they are based are based on the results of chronic toxicity tests on sensitive species.

Interferences are considered to be related to biotic or abiotic factors that could influence the SQGs derivation beyond the direct effects of specific contaminants. Because the SLCs are based entirely on benthic community data, they were considered to have the highest level of uncertainty. In contrast, residue-based guidelines are derived from direct analytical determination and are not subject to the same types of interferences. Uncertainty in the degree of standardization was evaluated on the basis of peer review. Approaches for determination of ESGs, ERLs and ERMs, and SEM-AVS have been published in the peer-reviewed literature and, hence, were assigned a low degree of uncertainty. In contrast, the mixture models (in the early stages of development with sediments), TRGs, and AETs had not been widely peer reviewed in the literature at the time of the evaluation. The results of various recent evaluations suggests that a lower level of uncertainty could be assigned to mixture models (e.g., USEPA 2000a).

SQGs were considered to be discriminatory if they could be used to correctly classify toxic and non-toxic samples. The ESGs and the ERLs and ERMs have been demonstrated to provide accurate tools for correctly predicting toxic and non-toxic responses in the field. In contrast, the SLCs have a poor ability to discriminate the range of adverse effects that could occur. Sediment samples with contaminant concentrations that exceed the AETs have a high probability of being toxic. However, the AETs may not reliably discriminate samples with lower levels of contamination with respect to their potential for adverse biological effects (i.e., false negatives). The factors that are considered to influence bioavailability are directly considered in the derivation of the ESGs, SLCs, SEM-AVS, and residue-based guidelines. Although other guidelines (i.e., ERLs and ERMs, AETs) are largely based on dry-weight concentrations, it is possible to refine the approaches to explicitly consider other normalization procedures.

Field validation was evaluated by Ingersoll *et al.* (1997) as an assessment of the predictability of the SQGs using a number of independent data sets (i.e., not used to derive the SQGs). Ingersoll *et al.* (1997) concluded that all of the SQGs listed in Table 2.5 were not adequately field validated. Subsequent to this analysis by Ingersoll *et al.* (1997), there have been numerous publications that have demonstrated the predictive ability of co-occurrence-based SQGs, such as ERLs and ERMs (e.g., Long *et al.* 1998b; Field *et al.* 1999; MacDonald *et al.* 2000a; 2000b; USEPA 2000b; Ingersoll *et al.* 2001a; 2001b).

In summary, Ingersoll *et al.* (1997) concluded that there is sufficient certainty associated with SQGs to recommend their use in assessments of sediment quality. In particular, ESGs, ERLs and ERMs, SEM-AVS, and residue-based SQGs generally have less uncertainty in their present applications than other guidelines. Although mixture models were generally considered to have somewhat higher levels of uncertainty, they address the critically important issue of the interaction of contaminants in complex mixtures. Importantly, a number or recent publications confirm that mixture models are essential for correctly predicting the presence and absence of sediment toxicity (MacDonald *et al.* 2000a; 2000b; Ingersoll *et al.* 2001b; USEPA 2000a). Toxicity identification evaluations (TIEs) and spiked-sediment exposures were recommended by Ingersoll *et al.* (1997) to help better establish cause and effect relationships between sediment chemistry and toxicity.

2.6 Interpretation of Data

Sediment chemistry data alone do not provide an adequate basis for assessing the hazards posed by sediment-associated contaminants to aquatic organisms or other receptors. Interpretive tools are also required to determine if sediment-associated contaminants are present at concentrations which could, potentially, impair the aquatic environment, wildlife, and/or human health. In this respect, the SQGs used in an assessment of sediment contamination need to provide a scientifically-defensible basis for evaluating the potential effects of sediment-associated contaminants on aquatic organisms, wildlife, and/or human health. Once the sediment chemistry data have been assembled, the quality and sufficiency of the data needs to be determined using explicitly defined evaluation criteria, such as those outlined in Appendix 4 of Volume III. If the sediment chemistry data do not meet the quality needed for the assessment, repeating certain components of the sampling program may be necessary.

The assessment of sediment chemistry data consists of three main steps (Figure 2.1). First, the measured contaminant concentrations at the sampling stations should be compared to regional background levels to determine if they are elevated relative to the background conditions (Appendix 2 of Volume III). Next, the concentrations of sediment-associated contaminants should be compared to applicable SQGs for the protection of aquatic life. Finally, the levels of contaminants in sediments should be compared to the bioaccumulation-based SQGs, including those for the protection of wildlife and the protection of human health. Problematic levels of contamination are indicated when sediment-associated contaminants are present at concentrations above one or more of the various SQGs and are present above background levels. However, the results of the sediment chemistry assessment should not be viewed in isolation. Instead, these results should be evaluated in conjunction with data on the other indicators assessed within measured at the assessment area (i.e., sediment toxicity, tissue chemistry, and community structure).

A variety of approaches have been used to determine if sediments exceed SQGs. For example, the number and/or magnitude of exceedances of individual SQGs has been used to classify sediment samples as toxic or non-toxic (i.e., MacDonald *et al.* 1996; USEPA 1996). Alternatively, procedures have been recently described for calculating combined effects of mixtures in sediment. Crane *et al.* (2000), USEPA (2000b), and Ingersoll *et al.* (2001b) described the relationship between mean probable effect concentration quotients (PEC-Qs) and the toxicity of whole sediments to amphipods and midges in short- and long-term exposure tests (see Appendix 3 of Volume III for a description of how PEC-Qs are calculated). Field *et al.* (1999; 2002) described a new procedure for evaluating matching marine sediment chemistry and toxicity data using logistic regression models. These models can be used to estimate the probability of observing an effect based on measured contaminant concentrations. Mixture models based on equilibrium partitioning have also been developed for assessing the toxicity of non-ionic organic compounds (Swartz *et al.* 1995; Di Toro and McGrath 2000) or metals (Ankley *et al.* 1996) in sediment.

The principal metrics for pore water chemistry are concentrations of contaminants in water. Targets for each of these metrics can be established from a variety of benchmarks for assessing water chemistry that have been published in the scientific literature. For example, numerical water quality criteria, such as those promulgated by the USEPA (1999), and site-specific water quality standards provide relevant tools to assessing pore water quality conditions. Alternatively, toxicity thresholds for pore water can be established using data available in the toxicological literature (i.e., median lethal concentrations or median effective concentrations; LC_{50}s or EC_{50}s) for receptors of concern at the site under consideration (Table 2.6). Such toxicity thresholds identify the concentrations of contaminants in water that are likely to cause acute and chronic toxicity to aquatic plants, amphipods and other aquatic invertebrates, and fish. USEPA (2000a) reported toxicity thresholds from 10-day water-only toxicity tests with the amphipod *Hyalella azteca*, the midge *Chironomus tentans*, and the oligochaete *Lumbriculus variegatus*, for a number of COPCs at contaminated sites.

Comparison of the concentration of a chemical in pore water to an LC_{50} or an EC_{50} for that chemical provides a means of determining if the concentration of that compound in the pore water was sufficient to cause direct toxicity to sediment-dwelling organisms (i.e., sufficient to cause sediment injury; Table 2.6). By dividing the pore water concentrations of each COPC in each sample by the reported LC_{50} concentration for that compound, it is possible to calculate a value that can be used to evaluate the overall toxicity of the sample. This value also provides a basis for reporting contaminant concentrations in terms of the number of toxic units. The number of toxic units of each compound can be summed to evaluate the combined toxic effect of chemicals with a similar mode of toxicity. Samples that contain ≥ 1 toxic units are likely to be toxic to sediment-dwelling organisms. See Ankley *et al.* (1996) for a description of an approach that was used to evaluate toxic units of metals in pore water samples.

Interpretation of sediment chemistry data relative to wildlife or human health necessitates the development of sediment quality targets that can be used to evaluate the extent to which these receptors are being protected. Such targets can be established by selecting appropriate SQGs for each bioaccumulative COPC at the site. Procedures for establishing sediment quality targets relative to direct human contact with contaminated sediments have been promulgated by United States Environmental Protection Agency and Agency for Toxic

Substances and Disease Registry. The bioaccumulation-based SQGs for the protection of wildlife or human health that were developed by the New York State Department of Environmental Conservation (NYSDEC 1999) and Washington State Department of Health (1995; 1996) provide a basis for establishing sediment quality targets relative to the protection of these receptors.

2.7 Recommendations

Sediment chemistry represents an essential indicator of sediment quality conditions in freshwater ecosystems. More specifically, sediment chemistry data are required to evaluate the nature, magnitude, and areal extent of sediment contamination. The following recommendations are offered to support the design and implementation of sediment quality assessments:

- The chemical analytes that are included in the sediment quality assessment program should include the COPCs that are identified based on the preliminary site investigation and the variables that support interpretation of the resultant data on the COPCs;
- Evaluations of the chemical composition of sediments should focus on determining the total concentrations of COPCs and SEM-AVS in whole-sediment samples. Analysis of other media types (e.g., pore water, elutriates) may also be conducted depending on the objectives of the investigation and the availability of resources;
- The benchmarks that are to be used in the sediment quality assessment should be identified in the data analysis plan, which is developed as part of the overall problem formulation process;
- Assuring the quality of sediment chemistry data is of fundamental importance to the integrity of the overall investigation. For this reason, it is important to design

and implement an effective QAPP for the program and include it as part of the sampling and analysis plan (SAP);

- The whole- sediment and pore water chemistry data that are generated during an investigation of sediment quality conditions should be evaluated relative to the project DQOs to determine which data are appropriate for use in the assessment (e.g., to determine if DQOs for accuracy, precision, and detection limits have been met);

- Numerical SQGs, such as consensus-based PECs and TECs (MacDonald *et al.* 2000a; 2000b; USEPA 2000b; Ingersoll *et al.* 2001a; 2001b) represent effective tools for assessing the potential effects of contaminated sediments on sediment-dwelling organisms (Tables 2.1 and 2.2). The potential effects of contaminated sediments on aquatic-dependent wildlife and human health can be evaluated using bioaccumulative-based SQGs, such as those that were derived by NYSDEC (1999);

- Toxicity thresholds for pore water provide useful tools for assessing the potential effects of contaminants on sediment-dwelling organisms (Table 2.6);

- Because contaminated sediments typically contain mixtures of COPCs, chemical mixture models [such as those developed by Swartz *et al.* (1995); Field *et al.* (1999; 2002); MacDonald *et al.* (2000b); USEPA (2000b); Ingersoll *et al.* (2001a; 2001b)] should be used to evaluate the effects of contaminated sediment on sediment-dwelling organisms;

- Whenever possible, decisions regarding the management of contaminated sediments should be made using a weight of evidence, which includes sediment chemistry and other relevant data. Nevertheless, the results of numerous evaluations of the predictive ability of SQGs indicate that sediment chemistry data can be used to accurately classify sediments as toxic or not toxic (i.e., typically with >75% correct classification using the results of whole-sediment toxicity tests). Therefore it is appropriate to make sediment management decisions using sediment chemistry data alone (i.e., with SQGs) at sites where the costs of further investigations are likely to approach or exceed the costs of sediment remediation; and,

- At sites where multiple indicators of sediment quality conditions are to be applied, sampling strategies should be developed and implemented that facilitate the collection of matching sediment chemistry and biological effects data (i.e., by preparing split samples for toxicity, chemistry, and benthos evaluations).

Chapter 3. Sediment Toxicity Testing

3.0 Introduction

Laboratory sediment toxicity tests can provide rapid and highly relevant information on the potential toxicity of contaminated sediments to benthic organisms. Acute (10- to 14-day exposures) and chronic (21- to 60-day exposures) toxicity tests have been developed to evaluate the biological significance of sediment contamination. Tests have been designed to assess the toxicity of whole sediments (solid phase), suspended sediments, elutriates, sediment extracts, or pore water. The organisms tested with these methods include microorganisms, algae, invertebrates, and fish. This chapter is intended to provide guidance on the selection of toxicity tests and interpretation of the associated results to support assessments of sediment quality conditions of contaminated sites.

3.1 Selection of Metrics and Targets for Sediment Toxicity

The objective of a sediment toxicity test is to determine whether contaminated sediments are harmful to benthic organisms (ASTM 2001a; USEPA 2000a). These tests can be used to measure interactive toxic effects of complex chemical mixtures in sediment. Furthermore, knowledge of specific pathways of interactions among sediments and test organisms is not necessary to conduct the tests. Sediment tests can be used to: (1) determine the relationship between toxic effects and bioavailability; (2) investigate interactions among chemicals; (3) compare the sensitivities of different organisms; (4) determine spatial and temporal distribution of contamination; (5) evaluate hazards of dredged material; (6) measure toxicity as part of product licensing or safety testing; (7) rank areas for clean up; and, (8) estimate the effectiveness of remediation or management practices.

The results of sediment toxicity tests can be used to assess the bioavailability of contaminants in field-collected sediments. The response of organisms exposed to field-collected sediments are often compared to the response of organisms exposed to a control and/or a reference sediment. The results of toxicity tests on sediments spiked with one or more chemicals can also be used to help establish cause and effect relationships between chemicals and biological responses. The results of toxicity tests with test materials spiked into sediments at different concentrations are often reported in terms of an LC_{50}, a median inhibition concentration (IC_{50}), a no observed effect concentration (NOEC), or a lowest observed effect concentration (LOEC; ASTM 2001a; USEPA 2000a).

The choice of a test organism has a major influence on the relevance, success, and interpretation of a test. As no one organism is best suited for all applications, considering the intended uses of the resultant data in the selection of toxicity tests is important. The following criteria were considered in the selection of the methods and species that were to be described in ASTM (2001a) and USEPA (2000a; Table 3.1). Ideally, a test organism should:

- Have a toxicological database demonstrating relative sensitivity and discrimination to a range of chemicals of concern in sediment;
- Have a database for inter-laboratory comparisons of procedures (for example, round-robin studies);
- Be in contact with sediment (e.g., water column vs. sediment-dwelling organisms);
- Be readily available through culture or from field collection;
- Be easily maintained in the laboratory;
- Be easily identified;
- Be ecologically or economically important;
- Have a broad geographical distribution, be indigenous (either present or historical) to the site being evaluated, or have a niche similar to organisms of concern (for example, similar feeding guild or behavior to the indigenous organisms);

- Be tolerant of a broad range of sediment physico-chemical characteristics (e.g., grain size); and,
- Be compatible with selected exposure methods and endpoints. The method should also be peer reviewed and confirmed with responses with natural populations of benthic organisms.

Of these criteria, a database demonstrating relative sensitivity to contaminants, contact with sediment, ease of culture in the laboratory, inter-laboratory comparisons, tolerance of varying sediment physico-chemical characteristics, and confirmation with responses of natural benthos populations were the primary criteria used for selecting the amphipod *Hyalella azteca* and the midge *Chironomus tentans* for describing test methods, as outlined by ASTM (2001a) and USEPA (2000a; Table 3.1). Procedures for conducting sediment tests with oligochaetes, mayflies, and other amphipods or midges are also outlined in ASTM (2001a) and in Environment Canada (1997b). However, USEPA (2000a) chose to not develop methods for conducting sediment toxicity tests with these additional organisms because they did not meet all the required selection criteria listed in Table 3.1. For both of the selected species (*Hyalella azteca* and *Chironomus tentans*), survival is the principal endpoint measured in acute toxicity tests (although growth is also commonly measured) and survival, growth, emergence (midges only) and/or reproduction are the principal endpoints measured in longer-term exposures.

USEPA (2000b) evaluated relative endpoint and organism sensitivity in a database developed from 92 published reports that included a total of 1657 field-collected samples with high-quality matching sediment toxicity and chemistry data. The database was comprised primarily of 10- to 14-day or 28- to 42-day toxicity tests with the amphipod *Hyalella azteca* (designated as the HA10 or HA28 tests) and 10- to 14-day toxicity tests with the midges *Chironomus tentans* or *Chironomus riparius* (designated as the CS10 test). Endpoints reported in these tests were primarily survival or growth. For each test and endpoint, the incidence of effects above and below various mean PEC quotients (mean quotients of 0.1, 0.5, 1.0, and 5.0) was determined. In general, the incidence of sediment toxicity increased consistently and markedly with increasing levels of sediment contamination. See Appendix 3 of Volume III for additional detail on calculation of mean PEC quotients.

A higher incidence of toxicity with increasing mean PEC-Q was observed in the HA28 test compared to the short-term HA10 or CS10 tests and may be due to the duration of the exposure or the sensitivity of growth in the longer HA28 test. A 50% incidence of toxicity in the HA28 test corresponds to a mean PEC-Q of 0.63 when survival or growth were used to classify a sample as toxic Figure 3.1 (USEPA 2000b). By comparison, a 50% incidence of toxicity is expected at a mean PEC-Q of 3.2 when survival alone was used to classify a sample as toxic in the HA28 test. In the CS10 test, a 50% incidence of toxicity is expected at a mean PEC-Q of 9.0 when survival alone was used to classify a sample as toxic or at a mean PEC-Q of 3.5 when survival or growth were used to classify a sample as toxic. In contrast, similar mean PEC-Qs resulted in a 50% incidence of toxicity in the HA10 test when survival alone (mean PEC-Q of 4.5) or when survival or growth (mean PEC-Q of 3.4) were used to classify a sample as toxic. The results of these analyses indicate that both the duration of the exposure and the endpoints measured can influence whether a sample is found to be toxic or not. The longer-term tests in which growth and survival are measured tended to be more sensitive than shorter-term tests, with acute to chronic ratios on the order of six indicated for *Hyalella azteca*. Based on these analyses, if only one of these tests were performed, it would be desirable to conduct chronic (i.e., 28- to 42-day) sediment toxicity tests with *Hyalella azteca* measuring survival and growth (as length) instead of 10- to 14-day tests with *Hyalella azteca*, *Chironomus tentans*, or *Chironomus riparius*.

Relative species sensitivity frequently varies among chemicals; consequently, both ASTM (2001a) and USEPA (2000a) recommend the use of a battery of tests to assess sediment quality, including organisms representing different trophic levels. However, testing multiple species with every sediment sample can be very costly. An alternate approach could be to perform a preliminary evaluation on a limited number of samples using a site with a battery of tests (i.e., see procedures for various species outlined in ASTM 2001a). This preliminary evaluation could be used to identify sensitive species or endpoints to include in a more detailed assessment at the site. The preliminary evaluation should include samples representing a gradient of contamination at the site of interest. This approach was taken by Kemble *et al.* (1994) in an assessment of the toxicity of metal-contaminated sediments in the

Clark Fork River in Montana. A battery of acute and chronic whole-sediment and pore water tests were conducted with samples collected from this site. The results of this investigation indicated that a 28-day whole-sediment toxicity test with *Hyalella azteca* measuring survival and growth (as length) was the most sensitive metric across a gradient of metal-contaminated stations at the site. The results of chronic toxicity test with *Hyalella azteca* were also predictive of effects observed on benthic community structure at the site (Canfield *et al.* 1994). Therefore, Kemble *et al.* (1994) recommended that future evaluations of sediment toxicity at the site should use chronic tests with *Hyalella azteca* rather than testing a suite of toxicity tests.

3.2 Availability of Standard Methods

Whole-sediment toxicity tests are the most relevant for assessing the effects of contaminants that are associated with bottom sediments. Standard methods have been developed for conducting whole-sediment toxicity tests with freshwater sediments by ASTM (2001a), Environment Canada (1997a; 1997b), and USEPA (2000a). The Organization of Economic Cooperation and Development (OECD) is in the process of developing standard methods for chronic sediment toxicity testing with midges. These methods can be used to assess the acute or chronic toxicity of sediment-associated contaminants on the amphipod, *Hyalella azteca*, the midges, *Chironomus tentans* and *Chironomus riparius*, the mayfly, *Hexagenia limbata*, and several other species of amphipods, cladocerans, and mollusks (Table 3.2). Standard methods have been recently described for conducting chronic whole-sediment toxicity tests with the amphipod *Hyalella azteca* and the midge *Chironomus tentans* (ASTM 2001a; USEPA 2000a). Endpoints measured in these chronic tests include effects on survival, growth, emergence (midge), and reproduction in 28- to 60-day exposures.

The procedures outlined in these standard methods can be modified to assess toxicity to other benthic invertebrate species that occur in freshwater environments. However, results of tests, even those with the same species, using procedures different from those described in the ASTM (2001a) and USEPA (2000a) may not be comparable and using these different procedures may alter the bioavailability of sediment-associated COPCs. Comparison of results obtained using modified versions of these procedures might provide useful information concerning new concepts and procedures for conducting sediment tests with aquatic organisms. If tests are conducted with procedures different from those described in ASTM (2001a) or in USEPA (2000a), additional tests (i.e., conducted on split sediment samples) are required to determine comparability of results.

Several endpoints are suggested to measure potential effects of contaminants in sediment, including survival, growth, behavior, or reproduction; however, survival of test organisms in 10-day exposures is the endpoint most commonly reported. Such short-term exposures, which only measure effects on survival, can be used to evaluate the effects associated with exposure to high levels of contamination in sediments, but may not be as relevant for assessing sediments with moderate levels of contamination (ASTM 2001a; USEPA 2000a). Long-term toxicity testing methods recently described in ASTM (2001a) and in USEPA (2000a) can be used to measure effects on reproduction, as well as long-term survival and growth. Reproduction is a key variable influencing the long-term sustainability of populations and has been shown to provide valuable and sensitive information in the assessment of sediment toxicity (ASTM 2001a; USEPA 2000a). Furthermore, as concerns have emerged regarding the environmental significance of chemicals that can act directly or indirectly on reproductive endpoints (e.g., endocrine disrupting compounds), the need for comprehensive reproductive toxicity tests has become increasingly important (SETAC 1999). Sub-lethal endpoints in sediment tests have also been shown to provide better estimates of responses of benthic communities to contaminants in the field (Hayward 2002).

The decision regarding the selection of short-term or long-term toxicity tests depends on the objectives of the assessment. In some instances, sufficient information may be gained by measuring growth in 10-day tests (i.e., for assessing highly contaminated sediments). However, longer term tests are needed to evaluate the effects associated with exposure to moderately contaminated sediments. Likewise, long-term tests are needed to directly assess effects on reproduction. Nevertheless, measurement of growth in these toxicity tests may serve as an indirect estimate of reproductive effects of contaminants associated with sediments (ASTM 2001a; USEPA 2000a).

Use of sub-lethal endpoints provides important information for assessing the ecological risks associated with exposure to contaminated sediments. As such, numerous regulatory programs require the use of sub-lethal endpoints in various decision-making processes (USEPA 2000a), including:

- Monitoring for compliance with water quality criteria (and state water quality standards);
- National Pollution Discharge Elimination System (NPDES) effluent monitoring (including chemical-specific limits and sub-lethal endpoints in toxicity tests);
- Federal Insecticide, Rodenticide and Fungicide Act (FIFRA) and the Toxic Substances Control Act (TSCA, tiered assessment includes several sub-lethal endpoints with fish and aquatic invertebrates);
- Superfund (Comprehensive Environmental Responses, Compensation and Liability Act; CERCLA);
- Organization of Economic Cooperation and Development (OECD, sub-lethal toxicity testing with fish and invertebrates);
- European Economic Community (EEC, sub-lethal toxicity testing with fish and invertebrates); and,
- The Paris Commission (behavioral endpoints).

ASTM (2001a) and USEPA (2000a) outline methods for measuring effects on reproduction in 42-day tests with *Hyalella azteca* or 60-day tests with *Chironomus tentans*. The results of water-only studies in chronic exposures to DDD, fluoranthene, or cadmium indicate that measures of reproduction are often more sensitive compared to measures of survival or growth for these species (Kemble *et al.* In preparation). However, these methods have not been applied routinely to assess the toxicity of field-collected sediments. Therefore, additional studies need to be conducted with field-collected sediments before these methods measuring reproductive endpoints are applied routinely to evaluate the toxicity of contaminated sediments.

ASTM (2001a) and USEPA (2000a) recommend additional research and methods development with standard methods for conducting sediment toxicity tests to:

- Evaluate additional test organisms;
- Further evaluate the use of formulated sediment;
- Refine sediment dilution procedures;
- Refine sediment TIE procedures;
- Refine sediment spiking procedures;
- Develop *in situ* toxicity tests to assess sediment toxicity and bioaccumulation under field conditions;
- Evaluate relative sensitivities of endpoints measured in tests;
- Develop methods for new species;
- Evaluate relationships between toxicity and bioaccumulation; and,
- Produce additional data on confirmation of responses in laboratory tests with natural populations of benthic organisms.

Some issues that may be considered in interpretation of test results are the subject of continuing research including the influence of feeding on contaminant bioavailability, nutritional requirements of the test organisms, and additional performance criteria for organism health.

In addition to whole-sediment toxicity tests, various procedures are available for assessing the potential for adverse effects on aquatic organisms due to the resuspension of sediments or partitioning of contaminants into pore water or into the water column. However, standard methods have not been developed for such methods. Perhaps the most frequently used of these is the bacterial luminescence test (Microtox; Schiewe *et al.* 1985; Burton and Stemmer 1988; Johnson and Long 1998) or cladoceran tests (Burton *et al.* 1996). Tests using algae, invertebrates, and fish also have been adapted to assess the toxicity of the suspended and/or aqueous phases, including pore water (ASTM 2001b). These exposures are typically conducted for 4 to 10 days, with survival measured as the primary endpoint. ASTM (2001a) and USEPA (2000a) describe procedures for isolating and handling pore water samples from whole-sediment samples.

3.3 Advantages and Disadvantages

Toxicity tests with aquatic organisms have a number of advantages that make them particularly relevant for evaluating the effects of contaminated sediments on aquatic organisms (Table 3.3; ASTM 2001a; USEPA 2000a). First, they provide quantitative information on sediment toxicity that provides a basis for discriminating between impacted and unimpacted sediment samples. In addition, standard methods have been established to support the generation of reliable data and minimize the effects of the physical characteristics of the sediments. The results of these tests are also ecologically- and socially-relevant because they commonly employ species which are familiar or important to area residents. Furthermore, studies conducted throughout freshwater environments in North America have demonstrated that aquatic organisms respond primarily to the contaminants in the sediments and pore water (i.e., not to physical factors or other variables; ASTM 2001a; USEPA 2000a). These characteristics make toxicity tests relevant for evaluating contaminant-related impacts in freshwater systems. Moreover, techniques for identifying the chemicals that are causing toxicity are being refined (i.e., TIE), which further support the identification of contaminants of concern (COCs; i.e., the substances that are causing or substantially contributing to sediment toxicity; USEPA 1991).

Toxicity tests also have several disadvantages which influence their application in sediment quality assessments (Table 3.3). For example, many of the tests that are currently used involve short-term exposures (i.e., 10-day) and, hence, may not be sensitive enough to detect sub-lethal effects on sensitive species. In addition, field collected sediments are manipulated before testing, which may affect their integrity and toxicity. Similarly, certain sediment phases (e.g., organic extracts, elutriates) may be less relevant for evaluating the *in situ* effects of toxic substances in sediments. Tests with field-collected samples may not discriminate effects of individual chemicals. Likewise, the ecological relevance of certain tests has not been fully established (e.g., Microtox; although it was not intended for this purpose but rather as an indicator of potential exposure). Importantly, certain test organisms may be more sensitive to certain classes of contaminants than others; therefore, it is desirable to use a suite of tests to cover the range of sensitivities exhibited by sediment-dwelling organisms in the field. See ASTM (2001a) and USEPA (2000a) for a more complete description of potential interferences associated with sediment toxicity tests.

Toxicity tests with fish also have several limitations which influence their application in sediment quality assessments. First, methods for assessing the toxicity of contaminated sediments to fish have not been standardized. In addition, toxicity tests with fish may be less sensitive than similar tests with freshwater invertebrates since fish derive more of their contaminant exposure from the overlying water (as opposed to exposure to pore water or during the processing of contaminated sediments). Furthermore, most of the tests that are currently available involve short-term exposures (i.e., 4- to 10-day) and, hence, may not be sensitive enough to detect sub-lethal effects on sensitive fish species. It is also difficult to obtain sufficient sample volumes to support testing with pore water. Finally, field collected sediments are manipulated prior to testing, which may affect their toxicity.

3.4 Evaluation of Data Quality

Use of performance-based methods have been recommended for use in sediment toxicity testing (ASTM 2001a; USEPA 2000a). Performance-based methods permit the use of appropriate methods that meet preestablished performance standards. For example, no one method must be used for culturing test organisms (ASTM 2001a; USEPA 2000a). However, having healthy test organisms of known quality and age for testing is critical to the success of the toxicity test. The performance-based criteria described in these methods allow laboratories to optimize culture methods and minimize effects of test organism health on the reliability and comparability of test results. A QAPP should be developed to address the experimental design and sampling procedures for the toxicity tests (Chapter 5 of Volume II).

Performance-based procedures are also established in ASTM (2001a), Environment Canada (1997a; 1997b), and USEPA (2000a) for establishing the acceptability of a toxicity test. For example, Table 3.4 from ASTM (2001a) and USEPA (2000a) outlines the method recommended for conducting chronic sediment toxicity tests with the amphipod *Hyalella azteca*, while Table 3.5 lists the test acceptability requirements for chronic sediment toxicity tests with *Hyalella azteca*. The primary requirements for meeting test acceptability include the age of organisms at the start of the exposure, minimum survival and growth of organisms at the end of the exposure in the control sediment, maintenance of water quality characteristics of the overlying water during the exposure, documentation of the quality of the cultures used to obtain organisms for testing, maintenance of the exposure system, and handling of sediments for testing (Table 3.5). ASTM (2001a) and USEPA (2000a) have provided specific definitions for the use of the terms "must" and "should" relative to test acceptability. "Must" is used to express an absolute requirement, that is, to state that a test ought to be designed to satisfy the specified conditions, unless the purpose of the test requires a different design. "Must" is used only in connection with the factors that relate directly to the acceptability of a test. "Should" is used to state that the specified condition is recommended and ought to be met if possible. Although the violation of one "should" is rarely a serious matter, violation of several will often render the results questionable. Additional Quality Assurance and Quality Control procedures for conducting sediment toxicity tests are outlined in ASTM (2001a), Environment Canada (1997a; 1997b), and USEPA (2000a).

3.5 Methodological Uncertainty

A review of uncertainty associated with the endpoints commonly used in sediment ecological risk assessments and approaches for addressing these sources of uncertainty was described in Ingersoll *et al.* (1997). The endpoints included in this evaluation included: toxicity tests (both the fraction tested and the endpoints selected); benthic invertebrate assessments; bioaccumulation assessments; sediment chemistry; and, sediment chemistry and SQGs.

A series of criteria were established by Ingersoll *et al.* (1997) to support consistent assessments of the uncertainty associated with each measurement endpoint. These evaluation criteria included: precision; ecological relevance; causality; sensitivity; interferences; standardization; discrimination; bioavailability; and, field validation (Table 3.6).

The results of these evaluations are presented in Tables 3.6 and 3.7. Uncertainty associated with lack of knowledge is indicated with an asterisk in these tables to differentiate from systematic uncertainty which can be rectified (methodologically) or quantified (sampling decisions and design). Uncertainty relative to laboratory toxicity tests was divided into two categories: Uncertainties related to the phase tested; and, uncertainties related to the selection of endpoints measured in toxicity tests (Ingersoll *et al.* 1997). A diverse array of exposure phases have been used in sediment toxicity tests. Six principal phases have been evaluated in toxicity tests (Table 3.6):

- Whole sediment using benthic invertebrates;
- Whole sediment using pelagic organisms;
- Organic extracts of whole sediment;
- Suspended solids
- Elutriates; and,
- Pore water isolated from whole sediment.

Whole-sediment toxicity tests were developed to evaluate the effects associated with exposure to in-place sediments. Toxicity tests with pore water samples isolated from sediment were developed for evaluating the potential *in situ* effects of contaminated sediment on aquatic organisms. Toxicity tests with organic extracts were developed to evaluate the effects of the maximum concentrations of organic contaminants associated with a sediment. Tests with elutriate samples and suspended solids measure the potential release of contaminants from sediment to the water column during disposal of dredged material or during sediment resuspension events.

Each of the six phases considered in sediment toxicity tests was evaluated in Ingersoll *et al.* (1997). The uncertainty associated with each phase is a function of inherent limitations of the test (e.g., testing of whole sediments has greater ecological significance than organic extracts) and the stage of development of the response as a toxicological endpoint (e.g., whole-sediment tests are much better developed than pore water tests). In Table 3.6, precision was evaluated in terms of the replicability the particular measurement. Ecological relevance was evaluated in terms of its linkage to the receptors which are to be protected. Causality was evaluated relative to the ability of the measure to determine the factors that adversely affect organisms exposed to contaminated sediments. Sensitivity was evaluated relative to the ability of the measure to identify sediments that have the potential to affect sensitive species in aquatic ecosystems. Interferences were evaluated related to biotic or abiotic factors which could influence the response of the measurement beyond the direct effects of specific contaminants. Standardization was evaluated in terms of the level of peer review and the publication of standard methods. Discrimination was evaluated based on whether or not a graded response could be identified. Bioavailability was evaluated relative to the ability of the measure to determine the fraction of contaminants in sediment that is readily available to organisms. Finally, field validation was evaluated relative to the extent to which the measure has been used to predict responses of benthic communities in the field.

Whole-sediment tests were considered to provide the most realistic phase for assessing organism response (Table 3.6). Because organic extracts may alter the bioavailability of sediment-associated contaminants, toxicity tests conducted using this phase were considered to have a relatively lower level of relevance. Similarly, elutriate and suspended solids tests are conducted using a phase which may artificially alter the availability of contaminants. In order to establish cause and effect relationships, it is necessary to link the toxicity test to appropriate measures of sediment chemistry, mixture toxicity models, spiked-sediment tests, and/or TIE procedures designed to help identifying specific compounds or classes of compounds responsible for toxicity. Ingersoll *et al.* (1997) provides a more complete summary of information presented in Table 3.6.

Uncertainties related to the selection of endpoints measured in toxicity tests focused on seven principal classes of response endpoints that are often measured in toxicity tests, including: survival; growth; reproduction; behavior; life tables; development; and, biomarkers (Table 3.7, Ingersoll *et al.* 1997). The uncertainties associated with each of the endpoints are a function of their inherent limitations (e.g., reproduction has greater ecological significance than biomarkers) and the stage of development of the response as a toxicological endpoint (e.g., acute lethality tests are much better developed than chronic reproductive tests).

The uncertainty associated with survival is less than that of the other endpoints used in sediment toxicity tests (Table 3.7). This is because mortality is an extreme response with obvious biological consequences. Also, a substantial body of literature concerning survival in sediment toxicity tests has been generated to date. Biomarkers have significant sources of uncertainty as sediment toxicological endpoints, especially with respect to ecological relevance and interferences by non-treatment factors. The continued development and application of more sensitive and ecologically relevant endpoints (e.g., chronic effects on growth and reproduction, life cycle tables) has the potential to produce superior measurement endpoints for use in assessment of contaminated sediments.

Toxicity tests, in and of themselves, are not useful for identifying the contaminants that are responsible for observed responses. Even linkage of test results to the list of chemicals measured during an exposure assessment might not provide all of the information needed to identify the potential causes of toxicity for a number of reasons including:

- Chemicals responsible for toxicity may not have been measured;
- The bioavailability of chemicals in either pore water or in whole sediment can be uncertain; and,
- Correlative techniques (i.e., comparison of responses to chemical concentrations) may be unable to deal with multiple contributions from complex mixtures.

Toxicity identification evaluation methods provide a useful approach for assessing toxicity contributions in sediment phases where unmeasured contaminants may be responsible for toxicity or where there are questions regarding bioavailability or mixture toxicity models (Ingersoll *et al.* 1997). The TIE methods consist of toxicity-based fractionation schemes that are capable of identifying toxicity due either to single compounds or to broad classes of contaminants with similar properties. Sediment TIEs have typically been conducted using pore water as the test phase; however, methods are being developed for testing whole sediments. Ingersoll *et al.* (1997) provides a more complete summary of the information presented in Table 3.7.

3.6 Interpretation of Data

For toxicity tests, the endpoints that are measured represent the primary metrics that are considered. Several methods have been used to establish targets for sediment toxicity tests. Most commonly, the response of test organisms (e.g., survival or growth) in test sediments are compared to responses in control sediments using a variety of statistical procedures. Samples in which the observed response of the test organism is significantly different from the control are designated as toxic. Similarly, the responses in test sediments can be compared to that in reference sediments, provided that the reference sediments are demonstrated to be appropriate (i.e., non-toxic, chemical concentrations below threshold effect-type SQGs; ASTM 2001a; Environment Canada 1997a; 1997b; USEPA 2000a; 2000b). For some toxicity tests (i.e., 10-day marine amphipod survival), power analyses have been used to identify minimum significant differences (MSD) from the control (i.e., the results of power analyses can be used to identify the response value that is always significantly different from the negative control, based on a specified alpha level; Thursby *et al.* 1997). Using this approach, test sediments are designated as toxic if the response of the test organism is significantly different from the control and the response rate exceeds the MSD from the control. Such MSDs have not yet been established for the freshwater toxicity tests that are commonly used in sediment quality assessments. Dilution series are often tested with pore water, elutriate, or organic extracts samples, with results typically reported as an LC_{50} (Carr *et al.* 1996).

Laboratory testing of sediment toxicity is an essential component of the sediment quality assessment process. At present, the nature and extent of available information on the effects of sediment-associated contaminants is such that there is often uncertainty associated with predictions of the biological significance of sediment-associated contaminants (i.e., most of the data available for field collected samples do not support the establishment of cause and effect relationships). Therefore, biological testing is required to provide reliable information regarding the toxicity of sediments (generally a suite of biological tests is desirable) and to confirm the results of the sediment chemistry assessment.

Further biological testing is required to support three distinct aspects of the sediment quality assessment process. First, biological testing may be required to assess the toxicity of sediments at stations where the concentrations of one or more contaminants is elevated above SQGs (e.g., PECs). Second, biological testing may be required to assess the toxicity of sediments that may contain unmeasured substances (i.e., based on the results of the preliminary site investigation). Third, biological effects data may be required to assess the site-specific applicability of the SQGs. In this respect, additional biological testing is required when the forms of the contaminants that are present may be less biologically available than those at other sites (i.e., the data used to support predictive ability evaluation of SQGs; USEPA 2000b).

The steps that should be used to assess sediment toxicity data are outlined in Figure 3.2. Once the sediment toxicity data have been assembled, the quality of the data needs to be evaluated in relation to the project DQOs (see Appendix 3 of Volume II). If the sediment toxicity data do not meet the quality needed for the assessment, it may be necessary to repeat certain components of the sampling and/or toxicity testing program.

The assessment of sediment toxicity data consists of two main steps (Figure 3.2). First, the results of the toxicity tests should be compared to the negative control data to determine if the sediments are significantly toxic. Next, the toxicity test results should be compared to data from appropriately selected reference stations. In this case, a reference sediment should be considered to be acceptable if it has been well-characterized and satisfies the criteria for negative controls (i.e., reference sediments should not be contaminated and reference results should not be significantly different from controls). Sediments that are found to be significantly toxic relative to control and reference sediments should be considered to be problematic. The results of the sediment toxicity assessment should be considered in conjunction with the results of the companion measures of other indicators of sediment quality, including sediment chemistry, benthic invertebrate community structure, and bioaccumulation, that are conducted at the site. ASTM (2001a) and USEPA (2000a) provide a description of procedures for conducting statistical analyses of data from toxicity tests.

3.7 Recommendations

The results of sediment toxicity tests provide important information for assessing the effects of contaminated sediments and aquatic organisms, including sediment-dwelling invertebrate species and fish. Based on the preceding evaluation of the applications of sediments toxicity test, the following recommendations are offered:

- Sediment toxicity testing should be included as an integral element of most sediment quality assessments;
- Because *in situ* communities of benthic invertebrates are exposed to contaminated sediments for extended periods of time, chronic toxicity tests are the most relevant for assessing effects on aquatic organisms;
- Due to their higher level of standardization and unequivocal relevance, whole-sediment toxicity tests should be preferentially included in sediment quality assessments; toxicity tests involving other media types (e.g., pore water) should be included as projects objectives and resources dictate;
- Although a wide variety of aquatic species may be tested, the amphipod, *Hyalella azteca*, and midge, *Chironomus tentans*, are the most highly recommended for most freshwater sediment quality assessments;
- Both lethal (i.e., survival) and sub-lethal (e.g., growth, reproduction, emergence) endpoints should be measured in sediment toxicity tests;
- Whenever possible, a suite of sediment toxicity tests should be used to assess sediment quality conditions;

- All sediments evaluated with toxicity tests should be characterized for at least: pH and ammonia of the pore water; and, organic carbon content (TOC), particle size distribution (percent sand, silt, clay), and percent water content of the sediment (ASTM 2001a; USEPA 2000a). Other analyses conducted on sediments can include: biological oxygen demand; chemical oxygen demand; cation exchange capacity; redox potential; total inorganic carbon; total volatile solids; AVS; metals; synthetic organic compounds; oil and grease; petroleum hydrocarbons; and, interstitial water analyses (ASTM 2001a; USEPA 2000a). The concentrations of other COPCs should also be measured, as identified on the PSI (Chapter 3 of Volume II);

- If direct comparisons are to be made, subsamples for toxicity testing should be collected from the same sample for analysis of sediment physical and chemical characterizations;

- Qualitative descriptions of the sediment should include color, texture, and the presence of petroleum sheens, macrophytes, or animals. Monitoring the odor of sediment samples should be avoided because of potential hazardous volatile chemicals;

- Following the selection of the most appropriate toxicity tests for the specific application, the test procedures and DQOs should be described in the project QAPP;

- The procedures for interpreting the sediment toxicity data should be described in the data analysis plan that is developed as part of the overall problem formulation process;

- The first step in the data interpretation process should involve evaluation of test acceptability (i.e., by comparing the results to the DQOs that were established in the QAPP);

- The results of sediment toxicity tests should be compared to those obtained for the negative control to evaluate test acceptability and/or to those obtained for appropriate reference sediment to assess the effect of contaminated sediment; and,

- Methods for testing caged organisms on site (i.e., *in situ* toxicity tests) are currently being developed by a variety of investigators (Crane and Maltby 1991; Veerasingham and Crane 1992; Seager *et al.* 1991; 1992; Maltby and Crane 1994; Crane, M. *et al.* 1995a; 1995b; 1996; 1999; 2000; Sarda and Burton 1995; Ireland *et al.* 1996; Chappie and Burton 1997; Olsen *et al.* 2001). These methods have been used to evaluate the acute toxicity of sediments in the field. However, additional methods development and standardization is needed before these methods are applied routinely to evaluate the toxicity of contaminated sediments.

Chapter 4. Benthic Invertebrate Community Assessment

4.0 Introduction

The structure of benthic invertebrate community structure represents an important indicator of sediment quality conditions. Such assessments are based on comparisons of community structure metrics, such as the diversity and abundance of benthic invertebrates, at test stations and appropriate reference stations (i.e., stations with similar depth, flow, sediment grain size, and TOC) and provide a means of assessing the relative impacts associated with exposure to sediments in the assessment area (USEPA 1992a; 1992b; 1994). Numerous studies have documented changes in the composition of benthic invertebrate communities resulting from sediment contamination (i.e., Rosenberg and Wiens 1976; Hilsenhoff 1982; 1987; Clements *et al.* 1992). However, many of these studies have examined the responses of benthic invertebrates in stony riffle areas of streams and rivers, and provide only limited information on the assessment of soft sediments (which typically accumulate elevated levels of contaminants; USEPA 1994). This chapter is intended to describe the existing procedures for assessing benthic invertebrate data as part of an overall assessment of sediment quality in depositional freshwater habitats.

4.1 Selection of Metrics and Targets for Benthic Invertebrates Community Structure

Benthic communities are assemblages of organisms that live in or on the bottom sediment. In most benthic community assessments, the primary objective is to determine the identity, abundance, and distribution of the species that are present (USEPA 1992a; 1992b; 1994). Because most benthic macroinvertebrates are relatively sedentary and are closely associated with the sedimentary environment, they tend to be sensitive to both short-term and long-term changes in habitat, sediment, and water quality conditions (Davis and Lathrop 1992). Therefore, data on the distribution and abundance of these species provide important information on the health of the aquatic ecosystem. As such, benthic invertebrate community structure represents an important ecosystem health indicator.

Assessments of benthic community structure have been used to describe reference conditions, to establish baseline conditions, and to evaluate the effects of natural and anthropogenic disturbances (Striplin *et al.* 1992). In terms of evaluating sediment quality, such assessments are focused on establishing relationships between various community structure metrics (e.g., species richness, total abundance, relative abundance of various taxonomic groups, and macroinvertebrate index of biotic integrity; mIBI) and measures of sediment quality (e.g., chemical concentrations, and organic content). Data from benthic community assessments have the potential to provide relevant information for identifying impacted sites and, with appropriate supporting data, the factors that are contributing to any adverse effects that are observed (USEPA 1992a; 1992b; 1994).

The International Joint Commission (IJC 1988) suggested that benthic community surveys should be the first assessment tool used to evaluate areas of the Great Lakes with suspected sediment contaminant problems. If no effects are demonstrated in an initial survey, IJC (1988) recommended no further assessment. However, the absence of benthic organisms in sediment does not necessarily indicate that contaminated sediment caused the observed response. Benthic invertebrate distributions may exhibit high spatial or temporal variability. Furthermore, short-term exposure to chemical (e.g., ammonia, dissolved oxygen) or physical (e.g., temperature, abrasion) factors can influence benthic invertebrate distribution and abundance, even in the absence of measurable levels of COPCs in sediment. Therefore, information on distribution of benthic invertebrates alone is not always indicative of ambient sediment quality conditions and is certainly not diagnostic of sediment contamination or sediment toxicity (USEPA 1992a; 1992b; 1994).

The objective of a benthic invertebrate community assessment is to determine whether sediment-associated COPCs may be contributing to a change in the distribution of benthic organisms in the field. These assessments can be used to measure interactive toxic effects of complex chemical mixtures in sediment. Furthermore, knowledge of specific pathways of interactions among sediments and test organisms is not necessary to conduct assessments of the benthic community. Assessments of the benthic invertebrate community can be used to:

- Determine the relationship between toxic effects and bioavailability;
- Investigate interactions among chemicals;
- Compare the sensitivities of different organisms;
- Determine spatial and temporal distribution of contamination;
- Rank areas for clean up; and,
- Evaluate the effectiveness of remediation or management practices.

The results of benthic community assessments can also be used to assess the bioavailability of contaminants in field-collected sediments. The response of organisms collected from test sites are often compared to the response of organisms collected from reference sites. Reynoldson *et al.* (1995; 1997) and MacDonald and Ingersoll (2000) describe procedures for assessing benthic invertebrate community structure of sediment quality conditions.

4.2 Availability of Standard Methods

Standard methods for collecting and processing sediment samples for evaluating benthic community characteristics have not been established by organizations such as the ASTM. This lack of standardization has resulted in the use of a wide variety of techniques to evaluate the effects of contaminated sediments on benthic invertebrate communities (Rosenberg and Resh 1993; USEPA 1992a; 1992b; 1994). These techniques can be classified into four general categories based on the level of organization that is considered (Ingersoll *et al.* 1997; Table 4.1), including:

- Individual (e.g., morphological changes, biomarkers);

- Population (e.g., abundance of keystone species; population size structure);
- Community structure (e.g., benthic index, multivariate analyses); and,
- Community function (e.g., energy transfer, functional groups).

All of the various measurement endpoints are evaluated based on departure from an expected or predicted condition (such as observations made at appropriate reference sites). Uncertainty in the application of these techniques stems from incomplete knowledge of the system (i.e., what represents normal conditions); systematic error in the method being used; and, the sampling scale that is selected (Ingersoll *et al.* 1997). One of the major limitations of these techniques is associated with the difficulty in relating the observed effect to specific environmental stressors (e.g., contaminants vs. low dissolved oxygen levels). For this reason, benthic invertebrate community structure has typically not been considered to be a central indicator of sediment quality conditions. However, such assessments may be conducted to provide ancillary information for further interpreting the sediment chemistry and toxicity data that are collected. Contingency tables have been developed for interpreting the results of sediment quality assessments that include multiple lines of evidence including benthic invertebrate assessments (Chapter 7 of Volume III). USEPA (1992a; 1992b; 1994) and ASTM (2001c) provide summaries of various procedures used to sample benthic invertebrates from sediments (i.e., grab samplers, artificial substrate samplers, dip nets; preservation and sorting of samples; taxonomic identification of samples).

4.3 Advantages and Disadvantages

Benthic invertebrate community assessments have a number of advantages that make them useful for evaluating the impacts of contaminated sediments on sediment-dwelling organisms (Table 4.2; USEPA 1992a; 1992b; 1994). First and foremost, the results of these assessments provide information that is directly relevant for evaluating benthic invertebrate community status (i.e., evaluating the *in situ* effects of contaminated sediments on the benthic community). In addition, procedures for conducting such assessments have been established that facilitate unbiased random sampling, support broad geographic coverage of the assessment area (including both contaminated and uncontaminated areas), and reduce variability in the results (i.e., by sampling under consistent hydrological and physical sediment conditions). Furthermore, the information generated is socially relevant (i.e., benthic species represent important food organisms for many sportfish species, such as walleye) and can be used to discriminate between sites that are degraded to various extents. The spatial and temporal distribution of benthic organisms may reflect the degree to which chemicals in sediments are bioavailable and toxic. Field surveys of invertebrates can provide an important component of sediment assessments for several reasons:

- Benthic invertebrates are abundant, relatively sedentary, easy to collect, and ubiquitous across a broad array of sediment types;
- Benthic organisms complete all or most of their life cycle in the aquatic environment, serving as continuous monitors of sediment quality; and,
- Assessment of indigenous populations may be useful for quantifying resource damage.

The usefulness of field studies with benthic invertebrates for assessing sediment contamination has been limited by several factors including:

- The composition of benthic communities has been difficult to relate to individual chemical concentrations;
- Benthic invertebrates respond to a variety of biotic and abiotic factors, in addition to contaminants
- Large numbers of samples are typically needed to address the high variance associated with distribution of benthos (USEPA 1992a; 1992b; 1994),
- A lack of standardized methods for collecting and processing samples; and,
- Inconsistencies in taxonomic identification of organisms.

Of primary concern, the information on benthic community structure can not be used alone to evaluate the cause of any impacts that are observed. While such communities certainly respond to chemical contamination in the sediment, they are also affected by a wide range of physical factors that are not directly related to sediment quality (e.g., low dissolved oxygen levels, grain size differences, nutritional quality of substrates, and water depth). In addition, benthic community composition exhibits significant spatial, short-term temporal, and seasonal variability; therefore, interpretation of the data relative to contaminant effects can be difficult. Care needs to be exercised to collect representative samples to minimize problems with data interpretation due to natural variation. For example, collection of samples should not be made after floods or other physical disturbances than may alter or remove benthic community assemblages (USEPA 1992a). The selection of reference sites can also influence the results of benthic community assessments. To complicate matters further, there is little agreement among benthic ecologists on which metrics are the most appropriate for evaluating the status of the community as a whole. Therefore, it is difficult to determine if information on individual organisms (e.g., morphological changes, biomarkers), populations of organisms (e.g., abundance of indicator species, population size structure), community structure (e.g., species richness, community indices), or community function (e.g., energy processing, presence of functional groups) should be used as indicators of benthic community status (Ingersoll *et al.* 1997).

4.4 Evaluation of Data Quality

Performance-based methods have been recommended for determining the acceptability of sediment toxicity tests (Chapter 3 of Volume III; ASTM 2001a; USEPA 2000a). Unfortunately, similar types of performance-based methods have not been established to determine the acceptability of benthic community data. A QAPP should be developed to address the experimental design and sampling procedures for the benthic community assessment (Chapter 5 and Appendix 3 of Volume II). The first step in conducting an evaluation of benthic invertebrate communities is the development of an appropriate experimental design (USEPA 1992a; 1992b; 1994). An inappropriate experimental design can be a major source of error in the resulting data. There are many factors to be considered when sampling contaminated sediments for benthic invertebrates that differ from the considerations required for sampling sediments for toxicity testing (Chapter 3 of Volume III). Benthic communities are strongly influenced by abiotic factors in the absence of contaminants, and in some cases, the effects of contaminants can be masked by effects due to abiotic factors. Important abiotic characteristics (i.e., sediment grain size, TOC, nutrient content, water quality, current velocity, and depth) at the site needs to be evaluated so that potential confounding effects of these characteristics can be accounted for when data are analyzed and interpreted. This holds true whether the intent of the project is to make comparisons between upstream and downstream areas, between different aquatic systems (different lakes or rivers), or between seasons.

When assessing benthic invertebrates for changes in community structure, it is critical to select appropriate reference sites (USEPA 1994; see Appendix 3 of Volume II). Ideally, reference sites should be unaffected or minimally affected by anthropogenic influences (ASTM 2001c). In addition to having low concentrations of COCs in sediment, the reference sites should also have physical and chemical characteristics of both water and sediment that are similar to the site under investigation to account for potential effects of these characteristics on benthic invertebrates.

Many studies have evaluated the number of replicate samples required to provide adequate assessments of benthic invertebrates (see USEPA 1992a; 1992b; 1994 for a listing of these publications). USEPA (1994) recommends that a sufficient number of replicate samples should be collected to achieve an among-sample coefficient of variation of less than 50%. Preliminary sampling at the sites of interest should be conducted to determine the number of replicates required to achieve this objective. Depending on the types of taxa collected, the methods used to collect samples may need to be modified to more effectively sample benthos at the sites of interest. The results of this preliminary study can also be used to determine the lowest practical level of taxonomic identification of the species at the sites of interest (USEPA 1992a). The data may not be normally distributed; therefore, transformation of data may need to be made to determine the appropriate number of replicates (USEPA 1992a). In addition, the variance may be different for the different endpoints evaluated (i.e., number of taxa vs. number of individuals). Previous studies have often collected three to five replicates per sampling station (USEPA 1992b; 1994). The decision to collect this number of replicates is often based on funding and personnel constraints that limit the processing of a large number of samples. Although the collection of a smaller number of replicates may not invalidate the benthic invertebrate data, such data should be interpreted with caution if the sites of interest are heterogeneous. USEPA (1992a; 1992b; 1994) include citations of several publications that more throughly address design of benthic invertebrate assessments.

4.5 Methodological Uncertainty

A review of uncertainty associated with the endpoints commonly measured in benthic invertebrate community assessments of sediment quality and approaches for addressing these sources of uncertainty was described in Ingersoll *et al.* (1997). A series of criteria were established by Ingersoll *et al.* (1997) to support consistent assessments of the uncertainty associated with each measurement endpoint. These evaluation criteria included: precision; ecological relevance; causality; sensitivity; interferences; standardization; discrimination; bioavailability; and, field validation.

The results of the evaluations of uncertainty associated with benthic community assessments are presented in Table 4.1. Uncertainty associated with lack of knowledge is indicated with an asterisk in this table to differentiate from systematic uncertainty which can be rectified (methodologically) or quantified (sampling decisions and design). Benthic invertebrates assessment methods were classified by Ingersoll *et al.* (1997) at different organizational scales, from the individual to the community level (Table 4.1). The types of endpoints included at these different organizational scales include:

- Individual (e.g., morphological changes, biomarkers);
- Population (e.g., indicator or keystone species abundance, population size structure and life history modifications);
- Community structure (e.g., indices, metrics, multivariate approaches); and,

- Community function (e.g., functional groups, energy transfer, size spectra).

Although community function was considered, there is little information on its use and application in sediment assessment. Therefore, the degree of uncertainty associated with its use is high because of lack of knowledge (Ingersoll *et al.* 1997).

The primary purpose of benthic invertebrate measurement metrics is to identify departure of the endpoint from either an expected or predicted condition, given normal variability in both time and space. Furthermore, these metrics should relate such a departure to a directional stressor. The precision of a benthic community assessment decreases as the scale of organization increases; thus, measurement of community metrics tends to be less precise than measurement of metrics relating to individual organisms. However, the uncertainty of measurements at the community level can be quantified and reduced by appropriate design and effort. Ingersoll *et al.* (1997) recommended that pilot studies be conducted to identify cost-effective benthic community metrics in relation to study objectives and available resources to reduce or quantify the uncertainty associated with problems of precision. Ecological relevance in Table 4.1 refers to the relationship between the measured endpoint and the benthic ecosystem. Accordingly, direct measures of the populations of organisms present have a higher certainty of being related to ecosystem than measurements at a finer organizational scale.

Measurements of benthic invertebrates provide little information on with which to identify the specific contaminants or stressors causing the response. Ingersoll *et al.* (1997) recommended that additional research be conducted, using controlled dose-response experiments, to evaluate the use benthic invertebrate data for identifying the toxic effects of specific contaminants in sediments (e.g., Hayward 2002). The response of benthic invertebrates may be sensitive to contaminants in sediment, but it is difficult to separate out effects due to interferences such as grain size, TOC, depth, and water quality characteristics of the overlying water at the site of interest. Additional standardization and field validation of methods used to assess and interpret benthic community data would improve the application of these approaches in sediment assessments, particularly in soft-bottom substrates where contaminants in sediments are of primary concern.

4.6 Interpretation of Data

A variety of metrics are directly relevant for assessing benthic invertebrate community structure (USEPA 1992a; 1992b; 1994; MacDonald and Ingersoll 2000). Domination of the benthic invertebrate community by pollution-tolerant species, such as worms (oligochaetes, particularly tubificid oligochaetes) and midges (chironomids), has been considered to be indicative of degraded conditions (i.e., for grab samples; MacDonald and Ingersoll 2000). The absence of more sensitive organisms, such as amphipods and EPT taxa (Ephemeroptera - mayflies, Plecoptera - stoneflies, and Tricoptera - caddisflies) has also been considered to provide strong evidence that benthic habitats and associated communities have been degraded. Additionally, mIBI scores were used to determine if benthic macroinvertebrate communities had been degraded relative to unimpacted sites (i.e., for artificial substrate samples; MacDonald and Ingersoll 2000). Information from studies on the colonization of benthic invertebrates on artificial substrates and from assessments of *in situ* benthic invertebrate community status can be used to assess benthic invertebrate community structure (USEPA 1992a; 1992b; 1994).

In general, sediment quality targets for the various metrics relating to benthic invertebrate community structure are established by assembling relevant information from relatively uncontaminated reference sites. For example, Ohio Environmental Protection Agency has established biocriteria applicable to the benthic community for a variety of ecoregions in the state using this reference site approach (OEPA 1988a; 1988b; 1989). Likewise, Simon *et al.* (2000) established a state-wide model for assessing benthic invertebrate community structure in Indiana using the mIBI, which provides a basis for establishing sediment quality targets. In this respect, Reynoldson *et al.* (1995) recommended that the normal range of benthic invertebrate community metrics be established using the 95% prediction limits; sediment quality targets could then be established as the upper and/or lower limits of the normal range for each metric (Reynoldson *et al.* 1997; Reynoldson and Day 1998; Reynoldson and Rodriguez 1999).

Benthic community assessments are required to support three distinct aspects of the sediment quality assessment process. First, benthic community assessments may be required to assess the effects of contaminated sediments at stations where the concentrations of one or more contaminants is elevated above threshold SQGs (e.g., PECs). Second, benthic community assessments may be required to assess the effect of sediments that could contain unmeasured substances. Third, benthic community assessment data may be required to assess the site-specific applicability of the SQGs. In this respect, additional data on sediment toxicity (Chapter 3 of Volume III) and on benthic community assessments may be needed when the forms of the contaminants that are present may be less biologically available than those at other sites (i.e., the data used to support predictive ability evaluation of SQGs; USEPA 2000a).

The steps that should be used to assess benthic invertebrate community status are outlined in Figure 4.1. Once benthic community data have been assembled, the quality of the data needs to be determined using criteria outlined in Section 4.4 of Volume III. If the benthic community data do not meet the quality needed for the assessment, it may be necessary to repeat certain components of the sampling and analysis program. The assessment of benthic community data consists primarily of comparing the response of individual metrics (i.e., number of taxa or an index) measured at test stations to those measured for appropriately selected reference stations (Figure 4.1). Test stations that are found to statistically differ from reference stations are classified as having a degraded community. These comparisons may be based on ANOVA, multivariate, or nonparametric statistical analyses (USEPA 1992a; 1992b; 1994).

4.7 Recommendations

The results of benthic invertebrate community assessments can provide useful information for evaluating the effects of contaminated sediments on sediment-dwelling organisms. Based on the preceding evaluation of the applications of benthic invertebrate assessments, the following recommendations are offered:

- Historically, sediment chemistry and toxicity data represent the primary elements of most routine sediment quality assessments. In some cases benthic invertebrate assessments have complemented these data by providing a basis for validating the results of such evaluations;

- The metrics that provide information on the status of the benthic invertebrate community (e.g., abundance of sensitive and tolerant taxa, species diversity, species richness, mIBI) are the most relevant for assessing sediment quality conditions;

- USEPA (1994) recommended a tiered approach for assessing benthic invertebrate communities. The first tier should include a qualitative preliminary survey of each study area to: (1) determine if community structure indicates alterations relative to reference conditions; (2) evaluate if there are differences in community structure across spatial gradients that may identify hot spots of contamination; (3) determine if taxa are represented by several orders of organisms or if the community is skewed toward a limited number of orders of organisms; and, (4) determine the number of replicate samples needed for the second tier of the assessment. Results from this first-tier assessment can be used to identify the best methods for sampling organisms at the sites of interest. The second tier should then include a quantitative survey that allows for a more robust statistical analyses of the various metrics chosen for the assessment.

- In order to interpret impacts on benthic invertebrates, it is critical to sample a number of reference stations that bracket the range in physical characteristics of the test stations. The physical characteristics that need to be considered when selecting a range of appropriate reference stations include sediment TOC, sediment grain size, water depth, water current, and water quality at the station;

- Benthic invertebrate community assessments should be designed to collect an adequate number of replicate samples from both reference and test sites to characterize within site variability;

- The procedures that are to be used to collect samples and to identify and count invertebrates should be documented in the QAPP;

- ASTM (2001a) and USEPA (2000a) recommend that all sediments evaluated with toxicity tests should be characterized for at least: pH and ammonia of the pore water; organic carbon content (TOC); particle size distribution (percent sand, silt, clay); and, percent water content. Other analyses on sediments can include: biological oxygen demand; chemical oxygen demand; cation exchange capacity; oxidation reduction potential; Eh; total inorganic carbon; total volatile solids; AVS; metals; synthetic organic compounds; oil and grease; petroleum hydrocarbons; and, interstitial water analyses (ASTM 2001a; USEPA 2000a). These physical and chemical characterizations of sediments are also relevant when collecting benthic community data at a site;

- Qualitative descriptions of the sediment may include color, texture, and presence of petroleum sheens, macrophytes, or animals. Monitoring the odor of sediment samples should be avoided because of potential hazardous volatile chemicals;

- The procedures for interpreting the results of the benthic invertebrate community assessments should be described in the data analysis plan that is developed as part of the overall problem formulation process;

- The first step in the data interpretation process should involve evaluation of data acceptability (i.e., based on the data quality objectives that were established in the QAPP);

- The results obtained for test sites should be compared with the results obtained for appropriately selected reference sites [i.e., uncontaminated sites which have similar physical (e.g., grain size, water depth), and chemical (e.g. dissolved oxygen) characteristics as the test sites];

- Models have been developed for use in predicting expected distributions of benthic invertebrates at stations in the absence of sediment contamination (Reynoldson *et al.* 1994). If these models are used, it is important to determine if the database used to develop the models is representative of the physical characteristics of the test stations being evaluated; and,

- Unlike the results of assessments conducted using sediment chemistry data, benthic invertebrate assessments alone should not be used to definitively determine sediment quality (USEPA 1992a). Again, the results of benthic invertebrate assessments should be considered in conjunction with the results of the companion measures of sediment chemistry, sediment toxicity, and bioaccumulation that are conducted at the assessment area (see Chapter 7 of Volume III).

Chapter 5. Bioaccumulation Assessment

5.0 Introduction

In aquatic ecosystems, many substances that occur at only trace levels in overlying water can accumulate to elevated levels in sediments. The same physical-chemical properties that cause these substances to accumulate in sediments (e.g., low aqueous solubilities, high K_{ow}), make chemicals such as PCBs, OC pesticides, and mercury prone to bioaccumulation. The accumulation of such substances in the tissues of sediment-dwelling organisms and subsequent bioconcentration and/or biomagnification in aquatic food webs can pose risks to a variety of ecological receptors, particularly those organisms that consume aquatic species. Bioaccumulation assessments are conducted to provide the information needed to assess the risks to aquatic-dependent wildlife and human health associated with exposure to bioaccumulative substances. This chapter is intended to describe the procedures for bioaccumulation assessments as part of integrated assessments, which represent important components of integrated assessments of sediment quality conditions.

5.1 Selection of Metrics and Targets for Bioaccumulation Assessment

Contaminated sediments represent important sources of the bioaccumulative substances that accumulate in aquatic food webs (Ingersoll *et al.* 1997). Because these contaminants can adversely affect aquatic-dependent wildlife species and/or human health, tissue chemistry represents an important ecosystem health indicator in sediment quality assessments (ASTM 2001d; USEPA 2000a). In general, the concentrations of contaminants in the tissues of sediment-dwelling organisms represent the primary metrics for tissue chemistry. As wildlife species typically consume the entire prey organism, whole body contaminant levels are the most relevant for assessing risks to wildlife. In contrast, the levels of contaminants in edible tissue represents the most important metrics for human health assessments. Assessments that are directed at evaluating contaminant residues in the tissues of benthic macroinvertebrates should focus on the bioaccumulative COPCs that are known or suspected to occur in sediments at the site under investigation. Typically, the COPCs that are considered in such assessments include: metals, methyl mercury, PAHs, PCBs, OC pesticides, chlorophenols, and/or PCDDs/PCDFs. However, this list should be refined based on the land and water use activities that have been documented in the vicinity of the site.

The selection of species for inclusion in assessments of bioaccumulation requires an understanding of the predator-prey relationships in the ecosystem under investigation. For example, the levels of contaminants in benthic macroinvertebrates are likely to be relevant when evaluating risks associated with dietary uptake of contaminants by bottom-feeding fish or sediment-probing birds. Conversely, emergent insects may be the primary focus of an investigation if swallows represent the primary receptor of concern. In cases where fish-eating birds and mammals represent the wildlife species of special concern, fish would be the primary species targeted in sampling and analytical programs. In this way, sampling programs can be tailored to answer the key risk questions that are being posed by the investigators. Bioaccumulation is not an appropriate assessment approach for contaminants that are metabolized or otherwise not accumulated in the tissues of the organism(s) being evaluated.

Ingersoll *et al.* (1997) identified four general approaches for conducting bioaccumulation assessments, including:

- A laboratory approach, which involves exposing organisms to sediment under controlled conditions;
- A field approach which involves collecting organisms from a study area;
- Assessment of food web transfer; and,
- Models to predict bioaccumulation processes.

The following sections briefly describe each of these approaches.

In the laboratory approach, individuals of a single species are exposed under controlled laboratory conditions to sediments collected from the study area being assessed (ASTM

2001d; USEPA 2000a). After an established period of exposure, the tissues of the organisms are analyzed for the COPCs. Bioaccumulation has occurred if the final concentration in tissues exceeds concentrations that were present before the exposure was started. This requires that individuals representative of initial conditions also be analyzed. This approach has been routinely applied in the assessment of contaminated sediments (ASTM 2001d; USEPA 2000a).

In the field approach, concentrations of COPCs in tissues are determined by collecting one or more species exposed to sediments at the study area being assessed. In addition, organisms representing various trophic levels may be collected and analyzed to determine tissue residue levels. These concentrations are compared to those that have been measured in the tissues of organisms collected from appropriately selected reference area(s). Two methods have been used to determine bioaccumulation in the field:

- Organisms resident at the area are collected *in situ* for analysis; or,
- Organisms are transplanted from another location (presumably with a history of little contaminant exposure) to the area of concern then re-collected, and tissues are analyzed after an established period of exposure.

These approaches have not been used routinely in the assessment of contaminated sediments (ASTM 2001d). In some cases, semipermeable membrane devices (SPMDs) are deployed in the field for specified time periods to simulate exposures of aquatic organisms to COPCs.

Models which describe bioaccumulation are relatively well developed for both organic and inorganic contaminants (Thomann 1989; Luoma and Fisher 1997; ASTM 2001d). Toxicokinetic models have a long history, as do simpler models of bioaccumulation processes. Site-specific models predict bioaccumulation on the basis of laboratory-determined characterization of biological processes in the species of interest and field-determined chemical measurements at the area of concern. Some uncertainties remain unresolved in most models and consensus does not exist about the appropriate model to apply for some (if not all) contaminants (Luoma and Fisher 1997).

Equilibrium models are commonly employed in risk assessment of bioaccumulation and are available for both organic and inorganic contaminants (Di Toro *et al.* 1991; Ankley *et al.* 1996). The models assume contaminant concentrations among all compartments of the environment are controlled by thermodynamics and, at least, approach equilibrium conditions. If thermodynamic equilibrium exists and if one route of uptake is known, or can be predicted, overall bioaccumulation is inferred. Recent applications use an extension of the equilibrium models, termed kinetic or pathway models (ASTM 2001d). These models incorporate geochemical principles and also address uncertainties in the assumptions of equilibrium. Kinetic models assume that routes of bioaccumulation are additive and must be determined independently. Kinetic models and equilibrium models may yield similar results if contaminant distributions and concentrations in an environment are at equilibrium (although not always), but can yield very different results where environmental compartments are not at equilibrium (e.g., if biological processes control concentrations, speciation, or phase partitioning of contaminants; Ingersoll *et al.* 1997).

Tissue residue guidelines for the protection of piscivorus wildlife species and/or human health represent the principal targets that are used to interpret the results of bioaccumulation assessments. However, a variety or risk-based procedures have also been developed to evaluate the results of such assessments. These tools can also be used to back-calculate to the concentrations of COPCs in sediment that will protect human health and ecological receptors.

5.2 Availability of Standard Methods

Standard methods have been developed for conducting whole-sediment bioaccumulation tests with a variety of test organisms, including the oligochaete *Lumbriculus variegatus* (ASTM 2001d; USEPA 2000a) and the amphipod *Diporeia* spp. (ASTM 2001d). The Organization for Economic Cooperation and Development (OECD) is in the process of developing standard methods for conducting sediment bioaccumulation tests with *Lumbriculus variegatus*. ASTM (2001d) also describes procedures for conducting sediment bioaccumulation tests with midges *(Chironomus tentans* and *Chironomus riparius*) and the amphipod (*Hyalella azteca*); however, *Lumbriculus variegatus* or *Diporeia* spp. are recommended in ASTM (2001d) for routine bioaccumulation testing with sediments.

The following criteria, which are outlined in Table 5.1, were used to select *Lumbriculus variegatus* for bioaccumulation method development (ASTM 2001d; USEPA 2000a):

- Ease of culture and handling;
- Known chemical exposure history;
- Adequate tissue mass for chemical analyses;
- Tolerance of a wide range of sediment physico-chemical characteristics;
- Low sensitivity to contaminants associated with sediment;
- Amenability to long-term exposures without feeding;
- Ability to accurately reflect concentrations of contaminants in field-exposed organisms (i.e., exposure is realistic); and,
- Data is available confirming the response of laboratory test organisms with natural benthic populations.

Thus far, extensive inter-laboratory testing has not been conducted with *Lumbriculus variegatus*. Other organisms that did not meet many of these selection criteria (i.e., as outlined in Table 5.1) included mollusks (valve closure), midges (short-life cycle), mayflies and *Diporeia* (difficult to culture), amphipods (*Hyalella azteca*; small tissue mass, too sensitive), cladocerans, and fish (not in direct contact with sediment).

Sediments for bioaccumulation testing may be either collected from the field or spiked with a range of concentrations of one or more COPCs. Recommendations are provided in ASTM (2001d) concerning procedures for meeting differing study objectives in sediment evaluations. These recommendations address the following: sediment physical and chemical measurements; test organism selection, collection, and maintenance; construction and maintenance of exposure systems; sampling methods and test durations; models that may be used to predict bioaccumulation; and statistical design of tests and analysis of test data.

The procedures outlined in these standard methods can be modified to assess bioaccumulation of contaminants in sediment by other benthic invertebrate species that occur in freshwater environments. However, the results of tests, even those with the same species, using procedures different from those described in the ASTM (2001d) and USEPA (2000a) may not be comparable, as using different procedures may alter the bioavailability of COPCs. If tests are conducted with procedures different from those described in ASTM (2001d) or in USEPA (2000a), additional tests are required to determine comparability of results. Comparison of results obtained using modified versions of these procedures might provide useful information concerning new concepts and procedures for conducting sediment tests with aquatic organisms.

Procedures described in these standard methods are designed to generate quantitative estimates of steady-state tissue residue levels, which are commonly used in ecological or human health risk assessments. Eighty percent of steady-state concentrations of sediment-associated contaminants is used as the general criterion for bioaccumulation tests. Because the results from a single or few species are often extrapolated to other species, the procedures are designed to maximize exposure to sediment-associated contaminants so that residues in untested species are not systematically underestimated. A 28-day bioaccumulation test with sediment-ingesting invertebrates, which are provided with no supplemental food, is recommended as the standard exposure scenario (ASTM 2001d; USEPA 2000a). Procedures for conducting long-term and kinetic tests are recommended for use when 80% of steady-state is unlikely to be obtained within 28 days or when more precise estimates of steady-state tissue residues are required (ASTM 2001d). The procedures are adaptable to shorter exposures and different feeding types. Exposures shorter than 28 days may be used to identify which compounds are bioavailable (that is, bioaccumulation potential) or for testing species that do not live for 28 days in the sediment (for example, certain species of midge such as *Chironomus tentans* or *Chironomus riparius*). Non-sediment-ingestors or species requiring supplementary food may be used if the objective is to determine uptake in these particular species due to their importance in ecological or human health risk assessments. However, the results obtained for such species should not be extrapolated to other species.

5.3 Advantages and Disadvantages

The strengths of using tissue chemistry data for evaluating the effects of contaminated sediments on sediment-dwelling organisms are similar to those that were cited for sediment chemistry data (Chapter 2 of Volume III; Table 5.2). These advantages include the availability of standard methods for quantifying contaminant concentrations in tissues, and of procedures for evaluating the accuracy and precision of the resultant data. Importantly, tissue chemistry data can be used to reliably identify the substances that are accumulating in the tissues of sediment-dwelling organisms and, as a result, causing or substantially contributing to sediment toxicity. Standard methods have also been developed for conducting bioaccumulation tests in the laboratory with sediments (ASTM 2001d; USEPA 2000a).

There are a number of factors that can limit the applicability of tissue chemistry data in sediment quality assessments. First, generation of high quality tissue chemistry data often requires a substantial mass of tissue to support analyses for the various COPCs. Collection of sufficient numbers of organisms to support such analyses can be challenging, particularly in highly contaminated sediments which typically have depauperate benthic communities. In addition, interpretation of such data is dependent on the availability of benchmarks that link tissue residue levels to adverse effects in sediment-dwelling organisms. The use of inappropriate analytical methods (i.e., with high reporting limits), the presence of interferences, and inadequate quality assurance practices can limit the utility of the resultant data. See ASTM (2001d) and USEPA (2000a) for a more complete description of potential interferences associated with conducting sediment bioaccumulation tests in the laboratory.

Tissue chemistry data provide important information for identifying the substances that are accumulating in biological tissues. However, these data cannot, by themselves, be used to assess risks or hazards to sediment-dwelling organisms. Interpretation of these data necessitates the establishment of targets that define the levels of COPC that are unlikely to adversely effect sediment-dwelling organisms. Bioaccumulated substances may cause an adverse effect on either the organism accumulating the material or an organism that consumes the contaminated tissue. While numerical TRGs are not yet available for assessing the direct effects of contaminant residues in benthic macroinvertebrates, Jarvinen and Ankley (1999) recently published a database that links tissue residues to effects on aquatic organisms. The United States Army Corps of Engineers has developed a similar database (Environmental Residue-Effects Database), which is available on the organization's website (http://www.wes.army.mil/el/ered/index.html). The information that is contained in these databases provides a basis for identifying toxicity thresholds (i.e., targets for tissue chemistry) for the various COPCs at the site under investigation. Subsequent comparison of field-collected tissue residue data to the published toxicity thresholds provides a basis for determining if bioaccumulative substances are present in the tissues at levels that are likely to be toxic to sediment-dwelling organisms.

The effects on aquatic-dependent wildlife associated with dietary exposure to tissue-borne contaminants are typically evaluated using numerical TRGs or toxicity reference values (TRVs) for tissues. In both cases, the measured concentrations of COPCs in the tissues of aquatic organisms are compared to the levels that have been established to protect piscivorus wildlife (TRGs; Newell *et al.* 1987) and/or the levels that are associated with specific types of adverse effects (TRVs; Sample and Opresko 1996). The potential for adverse effects on human health associated with the consumption of contaminated fish and/or invertebrate tissues can be evaluated using the Action Levels that have been established by the Food and Drug Administration (USEPA 1989). The availability of such benchmarks to support interpretation of the data represents an important advantage of the bioaccumulation approach.

5.4 Evaluation of Data Quality

The use of performance-based methods has been recommended for laboratory bioaccumulation testing (ASTM 2001d; USEPA 2000a). Performance-based methods permit the use of methods that meet preestablished performance standards (Chapter 3 of Volume III). The experimental design and sampling procedures for the bioaccumulation analyses should be documented in the project QAPP. Two primary issues related to quality of the data in bioaccumulation assessments include detection limits and replication. Detection limits for tissue analyses selected for the assessment should depend on the objectives of the study and the benchmarks for assessing potential effects (Section 5.6 of Volume III). ASTM (2001d) and USEPA (2000a) describe procedures for determining adequate tissue mass for the selected detection limits and minimum detectable differences among treatments. For example, ASTM (2001d) and USEPA (2000a) recommend a minimum of 1 g per replicate and preferably 5 g per replicate in bioaccumulation tests with the oligochaete *Lumbriculus variegatus*; five replicates per treatment were also recommended. Methods for achieving low detection limits for a variety of organic and inorganic compounds can be found in Ankley *et al.* (1992), Brunson *et al.* (1998), ASTM (2001d) and USEPA (2000a). Methods for achieving low detection limits for lipid analyses in small tissue samples can be found in Gardner *et al.* (1985), ASTM (2001d), and USEPA (2000a).

The decision to depurate the gut contents of organisms before chemical analysis is dependant on the objective of the study. If the objective of the study is to determine the total dose of contaminants in prey organisms that could be transferred to a predator, then test organisms should not be depurated before analyses of body burden. However, if the objective of the study is to determine a steady-state concentration of compounds in an organism, then organisms are typically depurated. See ASTM (2001d) and USEPA (2000a) for a discussion of approaches that can be used to estimate the contribution of contaminants in the gut to the overall body burden of contaminants in an organism.

Performance-based procedures have been established in ASTM (2001d) and USEPA (2000a) for establishing the acceptability of a laboratory bioaccumulation test. For example, Table 5.3 outlines a method for conducting 28-day sediment bioaccumulation exposures with the oligochaete *Lumbriculus variegatus*, while Table 5.4 lists the test acceptability requirements for conducting this test (ASTM 2001d; USEPA 2000a). The primary requirements for meeting test acceptability of organisms in this sediment exposure include behavior (i.e., organisms should not avoid the sediment) and toxicity (survival of organism should not be reduced relative to the control sediment), maintenance of water quality characteristics of the overlying water during the exposure, documentation on the quality of the cultures used to obtain organisms for testing (organisms at the start of the exposure should have low concentrations of COPCs), maintenance of the exposure system, and handling of sediments for testing (Table 5.4). Additional quality assurance and quality control procedures for conducting sediment toxicity tests are outlined in ASTM (2001d) and USEPA (2000a).

5.5 Methodological Uncertainty

In a review of uncertainty associated with endpoints commonly used in bioaccumulation assessments, Ingersoll *et al.* (1997) identified four general approaches for bioaccumulation assessments, including:

- A laboratory approach, which involves exposing organisms to sediment under controlled conditions;
- A field approach, which involves collecting organisms from a study area;
- Assessment of food web transfer; and,
- Models to predict bioaccumulation processes.

Each of these approaches was evaluated in Ingersoll *et al.* (1997) in relation to following major sources of uncertainty: precision, ecological relevance, causality, sensitivity, interference, standardization, discrimination, bioavailability, and field validation (Table 5.5). Precision was evaluated in terms of the replicability the particular measurement. Ecological relevance was evaluated in terms of its linkage to the receptors which are to be protected. Causality was evaluated relative to the ability of the measure to determine the factors that adversely affect organisms exposed to contaminated sediments. Sensitivity was evaluated relative to the ability of the measure to identify sediments that have the potential to affect sensitive species in aquatic ecosystems. Interferences were evaluated related to biotic or abiotic factors which could influence the response of the measurement beyond the direct effects of specific contaminants. Standardization was evaluated in terms of the level of peer review and publication of standard methods. Discrimination was evaluated in terms of whether or not a graded response could be identified. Bioavailability was evaluated relative to the ability of the measure to determine the fraction of contaminants in sediment readily available to organisms. Finally, field validation was established relative to how the measure has been used to predict responses of benthic communities in the field.

Variability is a common problem in bioaccumulation studies, and can lead to imprecise estimates of exposure, although standard methods for determining bioaccumulation describe procedures for avoiding extreme sources of uncertainty (ASTM 2001d; USEPA 2000a). Laboratory bioaccumulation tests are potentially the most precise of bioaccumulation approaches. However, their precision is directly dependent upon biological factors, such as the selection of appropriate test organisms. Number of individuals sampled, number of composites, life-stage, size of organisms, biases from analysis of gut content or surface contamination are examples of uncertainty associated with field approaches. Bioaccumulation models were ranked as imprecise because of the large knowledge gaps which remain in identifying values for model parameters (Table 5.5).

Ecological relevance includes both relevance to ecological change and relevance to human exposure pathways. A limitation to the bioaccumulation approach is its weak link to adverse ecological effects. Bioaccumulation does not mean an adverse effect is occurring. Organisms are capable of detoxifying, adapting to or otherwise surviving some dose of COPCs. Correlations between bioaccumulated contaminants and effects on sediment-dwelling organisms are also not as well established (Jarvinen and Ankley 1999). Collection of organisms exposed in the field, food web bioaccumulation estimates, and empirical and site-specific models provide direct determination of contaminant concentrations in aquatic resource (food) species and provide information for pathways of human exposure. Where tissue concentrations are directly determined in the food organism, there is little uncertainty about relevance. The precise human exposure pathway is predicted with less certainty if analyses of a surrogate species are used to estimate human exposures from a variety of species in an environment.

Causality describes the linkage between the source of the COPCs, exposure pathways, and the measured biological effect. Bioaccumulation data alone cannot provide information about whether the source of exposure was overlying water or sediment and cannot be used alone to evaluate effects of contaminants on aquatic organisms. Nevertheless, bioaccumulation data provide the strongest endpoints for drawing linkages to COPCs because it involves direct determinations of the concentrations of those substances in tissues. Bioaccumulation is a sensitive response because it measures exposure of an organism to relevant COPCs. However, bioaccumulation is not appropriate for determining exposures to ammonia or some metals, which are not bioaccumulated before exerting toxic effects. In addition, model results will be fraught with uncertainty about sensitivity until widely accepted input parameter values are established (Table 5.5).

Interferences can add uncertainties to bioaccumulation studies. Sediment characteristics are an important source of uncertainty in laboratory bioaccumulation studies because collection, transport, and deployment can change sediment characteristics from conditions in the field. It is possible that variability over small spatial scales interferes with or adds uncertainty to discrimination between areas on larger scales. Use of standard methods for field and laboratory bioaccumulation assessments can reduce uncertainty (ASTM 2001d; USEPA 2000a).

The ability of bioaccumulation to discriminate contamination gradients with low uncertainty is one of its advantages. Inherently, bioaccumulation is a highly quantitative approach for discriminating the risk of exposure to COPCs from a sediment. Bioaccumulation directly measures bioavailability in both laboratory and field studies. Some qualitative uncertainty in bioavailability (if it is defined as contaminants assimilated into tissues) can occur in determination of whole-tissue concentrations. Undigested gut content can be analyzed as part of the tissue burden and cause systematic uncertainties (upward bias) in estimates of bioavailability if contaminant concentrations in food are high compared to tissues (and if food mass in the gut is sufficiently great). Contaminants in gut content and on animal surfaces will be consumed by predators, so there is not a widespread consensus about the necessity of purging all undigested contamination from the gut of organisms. Some studies, especially with small organisms, have successfully related bioaccumulation obtained in the laboratory with field-collected sediments to residue concentrations observed in synoptically collected organisms from the field (Ankley *et al.* 1992; Brunson *et al.* 1998; Ingersoll *et al.* 2001c).

In summary, the principal use of bioaccumulation is to estimate the exposure or dose which organisms encounter in a sediment (Ingersoll *et al.* 1997). Bioaccumulation is not an appropriate assessment approach for contaminants which are metabolized or, for other reasons, not accumulated in the tissues of the organism(s) being evaluated. Another limitation of the bioaccumulation endpoint is its weak link to ecological effects. Bioaccumulation does not mean an adverse effect is occurring. The relevance of bioaccumulation stems mainly from its value in characterizing exposures and understanding the dose that an organism experiences. This can be especially valuable information if used to expand understanding of bioavailability or if exposures are complex in space or time (as is often the case) at the site of interest. Bioaccumulation can be a highly variable endpoint, but if established methods are followed and sample size is adequate, variability, imprecision and insensitivity can be controlled.

5.6 Interpretation of Data

Interpretation of tissue chemistry data relative to the potential for adverse effects on aquatic-dependent wildlife necessitates the establishment of targets that define tolerable levels of contaminants in the tissues of aquatic organisms. More specifically, such data may be compared to TRGs to determine if contaminants have accumulated in the tissues of aquatic organisms to such an extent that adverse effects on piscivorus wildlife species are likely to occur. Such TRGs for the protection of piscivorus wildlife have been developed by the New York State Department of Environmental Conservation (Newell *et al.* 1987). Toxicity thresholds for wildlife species have also been established to support interpretation of field and laboratory data (Sample and Opresko 1996).

The consumption of contaminated tissues represents the most important route of human exposure to bioaccumulative COPCs at sites with contaminated sediments. Fish consumption advisories are frequently established as a result of bioaccumulation of sediment-associated contaminants by fish (Beltman and Lipton 1998). For this reason, tissue chemistry represents an important ecosystem health indicator for assessing effects on human health. Application of this ecosystem health indicator necessitates the identification of appropriate metrics that can be used to evaluate the status of this indicator. A list of target analytes for biological tissues can be developed from the preliminary list of COPCs for the site (i.e., that is established using background information on the site) by identifying the substances that are likely to accumulate in biological tissues (e.g., mercury, certain PAHs, PCBs, organochlorine pesticides, PCDDs).

Evaluation of the actual hazards posed by bioaccumulative substances requires information on the levels of contaminants that are present in fish and shellfish tissues, on the weekly consumption of contaminated tissues by various sectors of the population, and on the toxicity of each contaminant to mammalian receptors. Alternatively, TRGs can be used, in conjunction with tissue residue data, to determine if existing concentrations of bioaccumulative substances pose a potential hazard to human consumers.

Interpretation of tissue chemistry data relative to the potential for adverse effects on human health necessitates the establishment of targets that define tolerable levels of contaminants in the tissues of aquatic organisms. In this context, numerical TRGs provide a basis for assessing sediment injury relative to human health. The Action Levels that have been established by the U.S. Food and Drug Administration (USEPA 1989) provide benchmarks for assessing the quality of fish tissues. Additionally, the presence of fish or wildlife consumption advisories provides direct evidence that the beneficial uses of the aquatic ecosystem have been compromised (i.e., the target for fish consumption advisories would be zero).

Information on levels of contaminants in aquatic biota and on bioaccumulation supports determination of the significance of contaminant levels in sediments relative to the direct toxic effects on these organisms or relative to protection of human health and the health of wildlife that consume these aquatic organisms. Equilibrium-partitioning models and kinetic models can also predict the accumulation of both organic and inorganic contaminants from sediment by aquatic organisms (ASTM 2001d).

Interpretation of tissue residue data is challenging for a number of reasons. While many aquatic organisms are sedentary (i.e., infaunal invertebrate species), others can be highly migratory (i.e., fish). For migratory species, it can be very difficult to establish where the exposure to bioaccumulative contaminants actually occurred. In addition, the concentrations of tissue-associated contaminants can vary depending on the trophic status, reproductive status, age, tissue sampled, and lipid content of the species under consideration, to name a few of the most important factors. Therefore, it is difficult to fully characterize the risks to wildlife and human health that are associated with the accumulation of contaminants in the food web.

Sediment characteristics, such as TOC, can have a major influence on the bioavailability of nonpolar compounds and increase the among-site variation in bioaccumulation (ASTM 2001d). Calculation of BSAFs can reduce this variability. Biota-sediment accumulation factors are calculated as the ratio of lipid-normalized tissue residue to organic carbon-normalized sediment contaminant concentration at steady state, with units of g-carbon/g-lipid. Normalizing tissue residues to tissue lipid concentrations reduces the variability in chemical concentrations among individuals of the same species and between species. These normalization procedures can be used to develop a simple thermodynamic-based bioaccumulation model for chemical uptake from sediment. The fundamental assumptions of this thermodynamic model are that the tissue concentration is controlled by the physical partitioning of the compound between sediment carbon and tissue lipids and that the organism and the environment approach thermodynamic equilibrium. The method assumes that lipids in different organisms and TOC in different sediments partition chemicals in similar manners. The key input parameter in the model is the BSAF, which predicts the lipid-normalized tissue residue when multiplied by the TOC-normalized sediment chemical concentration.

In theory, BSAFs should not vary with sediment type or among species. Based on the relationship between organic carbon partition coefficients (K_{oc}) and lipid-normalized concentrations in tissue, the maximum BSAF for neutral organic compounds has been calculated to be about 1.7 (ASTM 2001d). Measured BSAFs would be lower than this maximum if metabolism of the compound by the organism is rapid or the organism fails to reach steady-state body burdens due to limited exposure durations or kinetic limitations to accumulation (for example, steric hindrances to uptake and slow desorption from sediment particulates to interstitial water). Measured BSAFs could exceed the calculated thermodynamic maximum if there is active uptake of the chemical in the gut or if there is an increase in the gut fugacity of the chemical, driving the chemical from the gut into the body. The chemical fugacity in the gut could increase as the volume of food decreases during digestion or as a result of a reduction in lipids.

The steps that should be used to assess tissue chemistry data are outlined in Figure 5.1. Once tissue chemistry data have been assembled, the quality of the data needs to be determined using criteria outlined in Section 5.4 of Volume III and in ASTM (2001d) and USEPA (2000a). If the tissue chemistry data do not meet the quality needed for the assessment, it may be necessary to repeat certain components of the sampling program.

The measured concentrations of contaminants in biological tissues should be compared to regional background levels to determine if tissues contain elevated levels of contaminants (Figure 5.1). ASTM (2001d) and USEPA (2000a) provide a description of procedures for conducting statistical analyses of data from bioaccumulation assessments. Comparison of tissue chemistry data to published toxicity thresholds provides a basis for determining if bioaccumulative substances are present in the tissues of aquatic organisms at levels that are likely to be toxic to sediment-dwelling organisms or fish (e.g., Jarvinen and Ankley 1999). In addition, these data may be compared to numerical TRGs to determine if contaminants have accumulated in the tissues of aquatic organisms to such an extent that adverse effects on piscivorus wildlife species are likely to occur (Figure 5.1). Such TRGs for the protection of piscivorus wildlife have been developed by the New York State Department of Environmental Conservation (Newell *et al.* 1987). TRGs have also been developed for the protection of human health (USEPA 1989). The results of tissue residue chemistry should also be considered in conjunction with measures of sediment chemistry, sediment toxicity, and community status of benthic invertebrates and fish at the assessment area (Chapter 7 of Volume III).

5.7 Recommendations

The results of bioaccumulation assessments provide essential information for evaluating the uptake of bioaccumulative substances from contaminated sediments by sediment-dwelling and other aquatic organisms. In turn, this information provides a basis for evaluating the potential effects of bioaccumulative substances on aquatic-dependent wildlife and human health. The following recommendations are offered to support the design and interpretation of bioaccumulation assessments:

- Bioaccumulation assessments should be included as an integral element of freshwater sediment quality assessments that are conducted at sites that are known or suspected to contain bioaccumulative substances;

- The uptake of bioaccumulative substances from freshwater sediments should be evaluated using the results of 28-day bioaccumulation tests with the oligochaete, *Lumbriculus variegatus* (i.e., to support the determination of BSAFs and the prediction of levels in higher tropic level organisms). It is recommended that 28-day toxicity tests with the oligochaete *Lumbriculus variegatus* be conducted following procedures outlined in ASTM (2001d) and USEPA (2000a) and in Tables 5.3 and 5.4;

- The concentrations of bioaccumulative COPCs in test organisms exposed to control sediments should be determined at the beginning and end of the bioaccumulation test to support interpretation of the results of tests conducted using site sediments;

- The physical and chemical characteristics of sediments that are used in bioaccumulation tests should be determined, in accordance with the guidance provided in ASTM (2001d) and USEPA (2000a);

- The concentrations of bioaccumulative COPCs should be determined in sediment-dwelling organisms that are obtained from field-collected sediments to validate the results of laboratory bioaccumulation tests and to evaluate the potential for adverse effects on invertebrate-eating wildlife species (e.g., fish, sediment-probing birds);

- The concentrations of bioaccumulative substances in the tissues of aquatic organisms (fish and shellfish) from the site under investigation should be determined to evaluate the potential for adverse effects on aquatic-dependent wildlife and human health;

- A conceptual model of the site, including COPCs, potential exposure pathways, and receptors at risk, should be developed to guide the selection of species for bioaccumulation testing and tissue residue analysis;

- Following the selection of the most appropriate bioaccumulation test(s) for the specific application, the test procedures and DQOs should described in the project QAPP;

- The procedures for interpreting the results of the bioaccumulation tests and the tissue residue data for field-collected samples should be described in the data analysis plan that is developed as part of the overall problem formulation process;

- The first step in the data interpretation process should involve evaluation of test and data acceptability (i.e., by comparing the results to the DQOs that were established in the QAPP);

- The results of bioaccumulation tests should be compared to those obtained at the beginning of the test and/or those obtained for control sediments to evaluate the uptake of bioaccumulative COPCs;

- The results of bioaccumulation tests and the measured concentrations of bioaccumulative COPCs in aquatic organisms may be compared to toxicity reference values (TRVs) and/or TRGs to evaluate the potential for effects on aquatic-dependent wildlife and/or human health; and,

- Applications of exposure models and dose-response relationships provides a basis for refining the effects assessments that are conducted using the tissue residue data in conjunction with TRVs and TRGs.

The bioaccumulation of sediment-associated contaminants can best be determined by conducting laboratory bioaccumulation tests with sediments collected from the area of interest. Minimum physical and chemical characterization of sediment samples used in these bioaccumulation tests are outlined in Section 3.7 of Volume III dealing with sediment toxicity testing (see also ASTM 2001d and USEPA 2000a). In addition to laboratory bioaccumulation testing, it is also useful to collect organisms inhabiting sediments at the area of interest to determine the potential for food chain transfer of contaminants to upper trophic levels. It is critical to use analytical methods that have been previously demonstrated to meet the desired detection limits for tissue residues and lipids. It is also important to establish a minimum tissue mass per replicate needed for all of the required analyses before conducting an assessment of bioaccumulation with either field-collected or laboratory-exposed organisms.

Chapter 6. Fish Health and Fish Community Assessments

6.0 Introduction

Contaminated sediments have been demonstrated to be toxic to sediment-dwelling organisms and fish (MacDonald and Ingersoll 2000). More specifically, exposure to contaminated sediments can result in decreased survival, reduced growth, or impaired reproduction in benthic invertebrates and/or fish. Additionally, some contaminants in the sediments are taken up by organisms through bioaccumulation (Chapter 5 of Volume III). As a result, benthic organisms, fish, birds, and mammals can be adversely affected by contaminated sediments. This chapter describes procedures for assessing potential impacts of contaminated sediment on fish health and on the composition of fish communities.

6.1 Selecting Metrics and Targets in Fisheries Assessments

Data on fish health provides important information for determining if fish have been adversely affected by exposure to contaminated sediments. Fish health represents a relevant indicator of sediment quality conditions because fish that are exposed to contaminated sediment can exhibit impaired health. Health can be defined as the capacity of an organism to withstand stress (Schmitt *et al.* 2000). Hence, the more stressed (i.e., less healthy) an organism is, the less capacity it has to withstand further stress (Bayne *et al.* 1985). Assessments of fish health are intended to integrate the overall responses of an organism to environmental stresses, including exposure to toxic and bioaccumulative substances (Schmitt *et al.* 2000). Fish health represents a relevant indicator of sediment quality conditions as fish that are exposed to contaminated sediments can exhibit a variety of responses, some of which provide evidence of exposure to chemicals of potential concern (COPCs) and others which indicate that such exposures are adversely affecting the organism.

Investigators in the fish health field have utilized a number of metrics to assess exposure to toxic and bioaccumulative substances. For example, tissue chemistry data have been used extensively to quantify exposures to bioaccumulative substances, such as PCBs, PAHs, PCDDs/PCDFs, and OC pesticides (Table 6.1). In addition, a number of metrics, such as ethoxyresorufin-*O*-deethylase (EROD) activity in liver (responsive to PCBs, PAHs, and PCDDs/PCDFs), H4IIE assay results in whole fish (responsive to PCBs, PAHs, and PCDDs/PCDFs), sex steroid (estradiol and testosterone) levels in plasma (responsive to endocrine modulating substances), metallothein levels in liver and kidneys (response to metals), vitellogenin in plasma (response to endocrine modulating compounds), and macrophage aggregate analyses of spleen, kidney, and liver (responsive to PAHs and metals) have been used as evidence of exposure to various classes of contaminants (McCarthy and Shugart 1990; Schmitt *et al.* 2000; Table 6.1). While these metrics provide information on exposures to toxic and bioaccumulative substances, they do not provide direct information on the effects that are associated with such exposures. Therefore, more direct measures of the effects of contaminant exposures on fish heath are also needed in assessments of sediment quality conditions.

There are a number of metrics that can be used to provide information on the overall health of fish that have been exposed to elemental and organic chemicals. For example, histopathological examination of fish liver, gills, gonads, spleen, and kidney has been used to determine the frequency of lesions and tumors in fish (Malins *et. al.* 1985; Goyette *et al.* 1988; Payne *et al.* 1988). Somatic indices, such as the relative mass of gonads, spleen, and liver, have also been used as a measure of overall organism health (Grady *et al.* 1992). Furthermore, necropsy-based fish health assessments, which include visual examination of all tissues for external and internal abnormalities (e.g., deformities, fin erosion, lesions, tumors, parasites), can also be used to evaluate organism health (Nener *et al.* 1995; Antcliffe *et al.* 1997; Schmitt *et al.* 2000). These types of information on fish health status are important because impaired fish health can lead to increased rates of fish mortality and result in associated effects on fish populations.

Establishment of targets for fish health depends on the determination of normal conditions for the fish species that reside in the geographic area under consideration. In some areas (e.g., Indiana, Ohio), the incidence of deformities, fin erosion, lesions and tumors (i.e., DELT abnormalities) in fish have been determined for uncontaminated reference sites (Sobiech *et al.* 1994). As such, statistical comparisons can be made of the metric scores that are measured at contaminated site and the reference areas. In this way, it is possible to determine if fish health has been adversely affected at the site under investigation.

Exposure to toxic and bioaccumulative chemicals can adversely affect fish in several ways. First, exposure to chemical contaminants can cause behavioral abnormalities, increased incidence of disease, decreased fish health, impaired reproduction, and elevated levels of mortality. In addition, the presence of sediment-associated contaminants can impact the benthic invertebrate community and, thereby, reduce the abundance of preferred fish food organisms. As such, affected aquatic habitats may support only reduced populations of fish.

A variety of metrics can be used to assess the status of fish communities in freshwater ecosystems. Such metrics provide information on species composition (i.e., total number of species, types of species, percent sensitive species, and percent tolerant species), on trophic composition (i.e., percent omnivores, percent insectivores, and percent pioneer species), and on fish health (Karr and Chu 1997; 1999). Other metrics that have been used in various investigations include, species richness, total abundance, percent alien taxa, and trophic status (Karr and Chu 1999). Integration of these metrics into multimetric indices, such as the Index of Biotic Integrity (IBI), and the Index of Well-Being (IWB), provides a basis for evaluating the overall status of the fish community, rather than individual attributes of the community (Yoder and Rankin 1995; Karr and Chu 1999). In many areas, IBI and/or IWB scores have been determined for appropriately selected reference sites within the ecoregion under consideration (e.g., Indiana - Sobiech *et al.* 1994; Ohio - OEPA 1988a; 1988b; 1989; Florida - Griffith *et al.* 1994). In this way, the status of the fish community at a contaminated site can be compared with the community that would normally occur in areas with similar physical habitats, in the absence of chemical contamination. MacDonald and Ingersoll (2000) applied this approach to identify areas with the Indiana Harbor area of concern that had degraded fish communities.

6.2 Availability of Standard Methods

Standard methods for collecting and processing of fish samples have not been established by organizations such as the ASTM. Nevertheless, guidance has recently been published for evaluating fish health as part of the USGS biomonitoring of environmental status and trends (BEST) program (Schmitt *et al.* 2000). The BEST program has been designed to document temporal and spatial trends in fish health through the use of chemical and biological monitoring methods. Fish are normally selected for sampling based on:

- A high potential for exposure and response to COCs;
- Having a territory that overlaps the area being monitored; and,
- Being large and abundant enough to permit sampling.

Methods are outlined in the BEST protocols for measuring several metrics, including histopathology, EROD activity, lysozyme activity, macrophage aggregate analysis, H4IIE bioassay, vitellogenin, sex steroids, chemical analyses of whole fish, somatic indices, stable nitrogen isotopes, and necropsy-based fish health examination. See Table 6.1 for a brief description of each of these metrics. A general measure of overall organism health can be evaluated using metrics such as histopathology, lysozyme activity, or necropsy for internal or external abnormalities. Metrics such as H4IIE and EROD can be used to evaluate the potential for effects associated with specific classes of compounds such as PCBs, PAHs, or PCDDs/PCDFs.

6.3 Advantages and Disadvantages

Evaluation of fish health offers a number of advantages relative to the assessment of sediment quality conditions. First, fish are often keystone species in aquatic ecosystems (i.e., species that influence the structure and/or function of the ecosystem as a whole); therefore, data on fish health can provide relevant information for assessing the health of the ecosystem as a whole. In addition, human uses of aquatic ecosystems are often dependent on the availability and quality of sport and food fish. As impaired fish health can adversely affect such uses, fish health data can be used to assess the maintenance and restoration of the designated water uses. Importantly, certain contaminants that do not bioaccumulate to elevated levels in fish tissues can adversely affect their health (e.g., PAHs). Therefore, fish health assessments can provide relevant data for evaluating the effects of such contaminants (Malins *et al.* 1985; Payne *et al.* 1988).

While fish health assessment assessments can be highly relevant in evaluations of sediment quality conditions, there are several limitations that influence their applicability. First, assessments of fish health typically involve destructive sampling of large numbers of fish to support statistical comparisons between contaminated sites and reference areas, potentially impacting the populations of affected species. In addition, fish health can be affected by exposure to water-borne chemicals, as well as sediment-associated contaminants. Therefore, adverse effects cannot necessarily be attributed to contaminated sediments. Furthermore, fish can be migratory species that reside within the site under consideration for variable and unknown time periods. Hence it is difficult to fully determine the duration of exposure to contaminated sediments.

Many of the advantages that were cited for fish health assessments are also relevant to fish community assessments. That is, as keystone species in aquatic ecosystems, information on fish community status can provide valuable information on the health of the ecosystem as a whole. Additionally, changes in the composition of the fish community or the abundance of certain fish species have the potential to adversely affect the designated uses of a waterbody. Importantly, unlike fish health assessments, fish community assessments do not necessarily require destructive sampling and, hence, can be conducted without adversely affecting fish populations.

In spite of the advantages noted above, fish community assessments have a number of limitations that can influence their applicability in sediment quality investigations. First and foremost, fish communities can be affected by a variety of natural (e.g., flooding, drought) and anthropogenic (e.g., habitat alterations, fishing pressure, water-borne contamination, sediment-associated contamination) stressors. Additionally, fish are often not in direct contact with sediment; as such, it is challenging to determine the cause or causes of changes in the composition of the fish community. Furthermore, fish tend to be migratory species and, as such, the composition of fish communities can change on seasonal bases in response to natural factors, such as food supply, temperature changes, and reproductive status. Finally, the applicability of fish health and fish community data can be limited due to difficulties associated with obtaining sufficient samples to support statistical analysis of the data.

6.4 Evaluation of Data Quality

Performance-based methods have been recommended for determining the acceptability of sediment chemistry (Chapter 2of Volume III) or sediment toxicity tests (Chapter 3 of Volume III). Unfortunately, performance-based methods have not been established to determine the acceptability of fish health data or fish community data. The first step in conducting an evaluation of fish communities is the development of an appropriate experimental design. An inappropriate experimental design can be a major source of error in the resulting data. There are many factors to be considered when sampling fish that differ from the considerations required for sampling sediments (Chapter 2 of Volume III). Fish communities can be influenced by abiotic factors in the absence of contaminants, and in some cases, the effects of contaminants can be masked by effects due to these abiotic factors (Sobiech *et al.* 1994). Important abiotic characteristics (i.e., water quality, current velocity and depth, shade cover) at the site need to be evaluated so that potential confounding effects of these characteristics can be accounted for when data is analyzed and interpreted. This holds true whether the intent of the project is to make comparisons between upstream and downstream areas, between different aquatic systems (different lakes or rivers), or between seasons.

When assessing fish communities, it is critical to select appropriate reference sites. Ideally, reference sites should be unaffected or minimally affected by anthropogenic influences (ASTM 2001a; Appendix 3 of Volume II). In addition to having low concentrations of COCs in sediment, the reference sites should also have physical and chemical characteristics of both water and sediment that are similar to the study site to minimize the potential effects of these characteristics on fish communities. See Appendix 3 of Volume II for additional discussion of reference sites. The methods that are to be used in fish health and/or fish community assessments should be documented in the project QAPP.

6.5 Methodological Uncertainty

A review of uncertainty associated with endpoints measured in fish health or fish community assessments of sediment quality was not addressed in Ingersoll *et al.* (1997). Nevertheless the same criteria that were established by Ingersoll *et al.* (1997) can be used in this assessment to estimate uncertainty associated with measures of fish health and fish community structure in the assessment of sediment quality (Table 6.2) including: precision; ecological relevance; causality; sensitivity; interferences; standardization; discrimination; bioavailability; and, field validation.

The primary purpose of fish health or fish community metrics are to identify departure of the endpoint from either an expected or predicted condition, given natural variability in both time and space. Furthermore, these metrics should relate such a departure to a directional stressor. The precision of a fish community assessment was rated as moderate given movement of fish within the area of interest and the lack of direct contact of sediment by most fish. In contrast, fish health metrics were rated as relatively precise assuming that standardized methods are used to perform these evaluations. Ecological relevance in Table 6.2 refers to the relation of the measured endpoint to the fish community at the area of interest. Accordingly, direct measures of the fish health or fish communities have a high certainty of being related to ecosystem responses at the area of interest. However, some of the fish health endpoints provide an indication only of exposure.

Measurements of fish community structure provide limited information on specific contaminants or stressor causing the response. The response of fish may be to either contaminants in sediment or physical factors that interfere with interpretations of sediment quality, such as substrate, shade, flow, and water quality characteristics of the overlying water at the area of interest. In contrast, fish health metrics can be used to identify specific chemical stressors that may be causing adverse responses to organisms (e.g., EROD activity, lysozyme activity, macrophage aggregate analysis, H4IIE bioassay, vitellogenin, sex steroids). Neither fish health nor fish community metrics have been standardized through such organizations as ASTM; however, detail methods have been described for conducting these measures (OEPA 1988a; 1988b; 1989; Schmitt *et al.* 2000; USGS 2000). Methodological uncertainty relative to discrimination and bioavailability were both rated relatively high for fish community assessment given the difficulty in linking effects observed on fish to a specific location with contaminated sediments (Table 6.2). Because certain metrics used in fish health assessments respond to a specific class or classes of COPCs, the uncertainty associated with discrimination and bioavailability was considered to be lower. Both fish health and fish community metrics have been extensively field validated, but these assessments have not been routinely used to assess sediment quality.

6.6 Interpretation of Data

The steps that should be used to assess fish health data are outlined in Figure 6.1. Once fish health data have been assembled, the quality of the data needs to be determined using criteria outlined in Section 6.4 of Volume III. If these data do not meet the quality needed for the assessment, it may be necessary to repeat certain components of the sampling program.

Establishment of targets for fish health depends on determining normal conditions for the fish species that reside in the geographic area under consideration. For example, background conditions in terms of the incidence of DELT abnormalities in fish have been determined for areas in Indiana and Ohio (Sobiech *et al.* 1994). As such, fish health at test stations from these areas can be compared to the target for a geographic area being considered (Figure 6.1). If the incidence of adverse effects associated with fish health is not different from the geographic target, then fish health is unlikely to be adversely affected at the test station. However, if the incidence in abnormalities is higher than the geographic target, test stations are classified as having a degraded fish health.

As is the case for fish health, establishment of targets for the fish community necessitate determination of normal conditions for uncontaminated sites within the same ecoregion as the site under investigation. In Ohio, for example, data collected throughout the state have been used to generate IBI and IWB scores that denote exceptional, good, fair, poor, and very poor fish communities at three types of sites, including wading sites, boat sites, and headwater sites (OEPA 1988a; 1988b; 1989). Similarly, Indiana has calibrated the IBI for use in several ecoregions, thereby making it applicable for use in a number of areas within the state. In the absence of such benchmarks, normal conditions may be determined by selecting and sampling one or more reference sites that have similar habitat characteristics, but are unaffected by chemical contamination. The results of fish health assessments should be considered in conjunction with measures of fish community structure and results of companion assessments of sediment chemistry, sediment toxicity, and bioaccumulation that are conducted at the assessment area (see Chapter 7 of Volume III).

6.7 Recommendations

Fish health and fish community assessments provide useful ancillary information for evaluating exposure to, and the effects of, sediment-associated contaminants in freshwater ecosystems. Based on the forgoing evaluation of fish health and fish community assessments, the following recommendations are offered:

- Fish health assessments can be used to assess exposure of fish to certain classes of COPCs, including metals PAHs, PCBs, OC pesticides, and/or PCDDs/PCDFs;
- The metrics that provide the most direct information on exposure of fish to toxic and bioaccumulative COPCs include EROD, H4IIE, vitellogenin, and sex steroids;
- The metrics that provide the most direct information on the health of exposed fish include histopathology, lysozyme activity, somatic indices, and necropsy-based fish health assessments;

- The procedures that are to be used to assess fish health and fish community status should be documented in the QAPP;
- The procedures for interpreting the results of fish health and fish community assessments should be described in the data analysis plan that is developed as part of the overall problem formulation;
- The first step in the data interpretation process should involve evaluation of data acceptability (i.e., based on the DQOs that were established in the QAPP; and,
- The results obtained for test sites should be compared with the results obtained for appropriate reference sites [i.e., uncontaminated sites which have similar physical (e.g., grain size, water depth) and chemical (e.g., dissolved oxygen) characteristics as the test sites].

Chapter 7. Integration of Information on Multiple Indicators of Sediment Quality Conditions

7.0 Introduction

Sediment quality assessments are typically conducted to determine if sediments have become contaminated as a result of land or water use activities. When such contamination is indicated, the results of sediment quality assessments need to provide the information required to evaluate the nature, severity, and areal extent of sediment contamination. In turn, this information can be used to identify actual and probable use impairments at the assessment area. The purpose of this chapter is to describe procedures for interpreting the data that are generated for assessing effects on sediment-dwelling organisms, on aquatic dependent wildlife, or on human health (Chapter 5 of Volume I). Procedures for evaluating the quality of the data generated for specific indicators, such as sediment chemistry or sediment toxicity, are outlined in Chapters 2 to 6 of Volume III. Procedures for determining if specific targets for each of these individual indicators have been exceeded are also described in these earlier chapters. Importantly approaches for integrating data that are generated from multiple lines of evidence, including sediment chemistry, sediment toxicity, bioaccumulation, or responses of organisms in the field, are described in the following sections. A series of contingency tables (Tables 7.1 to 7.5) are presented which can be used to interpret impacts on aquatic life, wildlife, and human health, using a weight-of evidence approach.

7.1 Integration of Information on Multiple Indicators of Sediment Quality Conditions

While individual indicators of sediment quality each have an inherent level of uncertainty associated with their application, the uncertainty associated with an overall assessment of sediment contamination can be reduced by integrating information from each of these individual indicators. For example, sediment chemistry, sediment toxicity, and benthic community data can be used together in a sediment quality triad assessment to establish a weight-of-evidence linking contaminated sediments to adverse effects on sediment-dwelling organisms (Table 7.1). The integration of multiple tools using a weight-of-evidence approach has the potential to substantially reduce uncertainty associated with risk assessment of contaminated sediment and thereby improve management decisions (Long and Chapman 1985; Chapman 1992; Canfield *et al.* 1996).

The first step in the evaluation of sediment quality data should be to determine if individual indicators exceed the established targets. For example, the following questions should be addressed:

- Do the concentrations of COPCs in sediments exceed applicable SQGs (Figure 2.1)?
- Are sediments toxic relative to control and/or reference treatments (Figure 3.2)?
- Are communities of invertebrates in the field degraded relative to reference conditions (Figure 4.1)?
- Do the concentrations of COPCs in tissues exceed TRGs (Figure 5.1)?
- Is the health of fish compromised relative to reference conditions (Figure 6.1)?

The answers to these questions will help to establish if metrics associated with each of these individual indicators are adversely affected at the test stations relative to the reference stations. However, it is also important to determine the relationships among individual indicators measured at the assessment area. These relationships can be evaluated most directly by using scatter plots of the data to determine if there is correspondence between pairs of indicators and associated metrics measured on splits of individual samples collected from stations in the assessment area (e.g., sediment toxicity vs. sediment chemistry). Alternatively, the scatter plots can be used to evaluate broader trends across geographic reaches within the assessment area (e.g., fish community status or fish health vs. sediment chemistry). Comparisons of fish community status or tissue chemistry of fish are often made across multiple stations sampled for sediment chemistry to account for the movements of fish within the assessment area.

Statistical regression analyses can be used to determine if there are significant relationships between pairs of indicators and associated metrics. For example, Figure 7.1 illustrates the relationship between sediment chemistry (as a function of mean PEC-Qs) and sediment toxicity (as a function of toxicity to *Hyalella azteca* in 10-day sediment tests). Similarly, relationships between metrics for a particular indicator can also be evaluated using scatter plots. Figure 7.2 illustrates the relationship between two metrics for sediment chemistry: SEM normalized to AVS (i.e., SEM-AVS) and toxic units of metals measured in pore water from these same samples. The results of these types of analyses can be used to establish concordance among various indicators (i.e., high chemistry and toxic, low chemistry and not toxic). Additionally, these analyses can help to establish the rate of false positives (i.e., high chemistry and not toxic) or false negatives (i.e., low chemistry and toxic) among various indicators.

The following sections describe procedures for using contingency tables in an expanded version of the sediment quality triad approach to incorporates measures of bioaccumulation with the traditional measures of sediment quality (MacDonald 1998). Specifically, integration of data from sediment chemistry, sediment toxicity, community status, and/or tissue chemistry provides important information for assessing sediment quality conditions. The contingency tables presented in Tables 7.1 to 7.5 provide a means of interpreting the data generated from multiple indicators of sediment quality using a weight-of-evidence approach. The results of these analyses can be used to estimate the likelihood of impacts of sediment contamination on aquatic life (sediment-dwelling organisms), wildlife (vertebrates), or human health.

7.1.1 Integration of Information on Multiple Indicators for Assessing Impacts on Sediment-Dwelling Organisms and Other Receptors

Historically, the sediment quality triad is the approach that has been used most frequently to evaluate the concordance between measures of sediment chemistry, sediment toxicity, and benthic community structure in the assessment of impacts of on sediment-dwelling organisms. The continency table presented in Table 7.1 presents eight possible outcomes based on the correspondence among these three indicators of sediment quality. Alternatively, broader assessments of sediment quality conditions can be conducted by also considering the potential for bioaccumulation. There are 16 possible outcomes when four individual indicators of sediment quality are evaluated (sediment chemistry, sediment toxicity, benthic community surveys and tissue chemistry; Table 7.2) providing a basis for assessing effects on sediment-dwelling organisms, aquatic-dependent wildlife, and human health. Frequently, there may only be two indicators of sediment quality reported for a particular site assessment (i.e., chemistry and toxicity), which would result in a contingency table with four possible outcomes (Table 7.3).

In each of these contingency tables, a "+" or "-" within in a column and row designates that the indicator for a particular sample (or station) is classified as being adversely affected "+" or not "-" relative to the established target. Multiple metrics can be used in classifying an individual indicator as impacted or not impacted. For example, multiple sediment toxicity tests or multiple measures of sediment chemistry may be reported for splits of the same sample collected from a station. MacDonald and Ingersoll (2000) classified a sample as toxic if one or more of the tests on a sample exceeded the target for toxicity relative to control or reference sediments. Similarly, a sample was designated as impacted if one or more measures of sediment chemistry exceeded established targets for selected SQGs. Alternatively, Canfield *et al.* (1994; 1996) described a procedure for ranking multiple metrics for a particular indicator to designate a sample (or station) as impacted. Menzie *et al.* (1996) describe a procedure for assigning weighting factors when ranking multiple metrics in an ecological risk assessment. Carr *et al.* (2000) described a procedure for using principal component analyses to classify indicators of sediment quality as impacted relative to reference conditions.

Concordance among the various indicators of sediment quality measured on the same sample generate a high level of confidence that the sample is being correctly classified as impacted or not impacted. For example, if each of the four indicators of sediment quality were designated as adversely affected (line 1 in Table 7.2), it would be highly likely that the station is impacted due to contaminant-induced degradation in the field resulting in direct toxicity and bioaccumulation. Similarly, if all of the indicators except for bioaccumulation indicated that a station is impacted (line 9 in Table 7.2), it is highly likely that the station is being adversely affected by the toxic substances present in contaminated sediments. In this case, however, bioaccumulative substances are not contributing to use impairment. Alternatively, if each of these four indicators of sediment quality were designated as not adversely affected (line 10 in Table 7.2), it would be highly unlikely that the station is being impacted. There may be stations where the individual indicators are not in concordance. For example, there may be no indication of effects based on sediment chemistry, toxicity, or benthic community structure, but bioaccumulation is occurring, based on exceedances of tissue chemistry targets (line 2 in Table 7.2). In this instance, it is unlikely that contaminants would be directly toxic to organisms at the station. However, adverse effects on aquatic-dependent wildlife and/or human health could be occurring.

There may be instances where sediment toxicity, benthic community structure, or tissue chemistry identify a station as impacted, but sediment chemistry is not elevated (i.e., lines 4, 7, or 15 in Table 7.2). In these instances, the station may be impacted as a result of unmeasured contaminants contributing to the toxicity. In other instances, there may be impacts identified with sediment chemistry and toxicity, but community structure is not impacted (i.e., lines 6 or 14 in Table 7.2). This situation may be the result of spatial variability of contaminants in the field that is not identified with composited samples used to measure chemistry and toxicity. Impacts on benthos in the field without corresponding impacts identified with sediment chemistry or toxicity may also result from spatial (or temporal) variability of contaminants in the field (i.e., lines 5 and 13 in Table 7.2). However, effects on organisms in the field also may reflect differences in habitat or other physical factors (i.e., low dissolved oxygen) rather than reflecting responses to contaminants (line 13 in Table 7.2). The presence of elevated levels of bioaccumulative contaminants in tissues indicates the potential for adverse effects on aquatic-dependent wildlife and/or human health.

Sediment may not be toxic in laboratory tests, but there may be elevated levels of contaminants, bioaccumulation, or evidence of altered benthic community structure (lines 3, 8, and 16 in Table 7.2). In these instances, the toxicity tests may not be sensitive enough to detect toxicity in the laboratory or chemicals in the sediment may not be directly toxic to organisms in the field. Sediment may also have elevated levels of contaminants without any other indication of sediment impacts (line 11 in Table 7.2). In these instances, there may be contaminants that are not bioavailable in the sediments. Alternatively, the target SQGs may be too low. For example, if the targets for sediment chemistry were based on exceedances of threshold-type SQGs (i.e., effects range-lows [ERLs] or threshold effect levels [TELs]), then there may be a high rate of false positives (SQG exceeded and non-toxic sample). Finally, there may be instances where sediments are identified as toxic in laboratory tests without any other indication of sediment contamination (line 12 in Table 7.2). In these instances, there may be unmeasured chemicals contributing to the toxicity. Alternatively, the sediment toxicity test may be responding to an abiotic characteristic of the sediments that is out of the tolerance range of the test organism (i.e., TOC influencing the growth of midges; ASTM 2001a).

The simplest contingency table, where only two indicators of sediment quality have been measured at the sampling stations, is presented in Table 7.3. In this example, sediment chemistry and sediment toxicity are being compared and there are only four possible outcomes. A station could be identified as impacted or not impacted due to toxicity and chemistry exceeding the established targets (lines 1 and 2 in Table 7.3). Elevated chemistry with no toxicity may be classified as a false positive (line 3 in Table 7.3). In this instance, the target thresholds for sediment chemistry may be set too low. Alternatively, the toxicity test may not have been sensitive enough to detect the elevated chemicals in the sample. A sample identified as toxic without elevated chemistry would be classified as a false negative (line 4 in Table 7.3). Perhaps the toxicity tests was responding to abiotic characteristics of the sediment (i.e., TOC or ammonia). Alternatively, there may be unmeasured chemicals contributing to the toxicity. Clearly, the use of only two indicators limits the overall interpretation sediment quality at a the assessment area. Ideally, sediment chemistry, sediment toxicity, benthic community structure, and tissue chemistry, would be measured at all stations to provide a more robust evaluation of sediment quality (Table 7.2).

Contingency tables are useful for determining concordance among various indicators of sediment quality. Canfield *et al.* (1998) used a the contingency table similar to Table 7.1 to determine the percentage of stations in an assessment area classified in each of the eight possible outcomes. A second approach for evaluating concordance among individual indicators of sediment quality would be to plot the data on a map (Figure 7.3). Data for individual indicators in these tri-axial graphs were arithmetically scored proportionally between 1 and 100 (i.e., 1 is indicative of the lowest concentration, least toxic, or most robust benthic community observed and 100 is the most impacted; Canfield *et al.* 1994). More than one metric can be used for a particular indicator by scoring each individual variable, summing these scores across the individual metrics, and re-scoring the sum of the combined scores between 1 and 100. The results of these analyses can then be plotted on tri-axial graphs when three indicators are being evaluated (Figure 7.3). Alternatively, these plots could include multiple axes if additional indicators are being evaluated (i.e., quad-axial graphs for the contingency table presented in Table 7.2). These plots are useful for evaluating general trends among stations at the assessment area. However, symmetry among the individual indicators in these plots does not always represent concordance among the indicators. There may be instances where a relatively low score for sediment chemistry or toxicity is identified as impacted relative to the target, whereas a higher score for benthic community would be needed to identify a station as impacted relative to the corresponding target (Canfield *et al.* 1996).

Carr *et al.* (2000) presented an alternative procedure for plotting the results of a sediment quality triad investigation on a map of the study area. Color-coded pie diagrams for each station were subdivided into three sections and each section was used to classify chemistry, toxicity, or benthic community as indicating high (green), medium (yellow), or low (red) sediment quality.

An example application of the sediment quality triad assessment of sediment quality was presented in a series of papers by Canfield *et al.* (1994; 1996; 1998). Sediment toxicity, chemistry, and benthic community structure were measured at stations located in the following areas:

- Three Great Lakes AOCs (Buffalo River, NY; Indiana Harbor, IN; Saginaw River, MI);
- The upper Mississippi River; and,
- The Clark Fork River located in Montana.

The results of the benthic invertebrate community assessments were compared to the sediment chemistry and toxicity data for each site. Good concordance was evident between measures of laboratory toxicity (28-day sediment exposures with *Hyalella azteca*, which measured effects on survival, growth, and sexual maturation), sediment contamination, and benthic invertebrate community composition in highly contaminated samples. However, in moderately contaminated samples, less concordance was observed between the composition of the benthic community and either laboratory toxicity test results or sediment contaminant concentrations. Laboratory sediment toxicity tests which measured sub-lethal endpoints better identified chemical contamination in sediments compared to many of the commonly used measures of benthic invertebrate community composition. One explanation for this is that the benthic community attributes may reflect other factors, such as habitat alterations, in addition to responding to contaminants. Canfield *et al.* (1994; 1996; 1998) concluded that there is a need to better evaluate non-contaminant factors (i.e., TOC, grain size, water depth, habitat alteration) in order to better interpret the response of benthic invertebrates to sediment contamination.

Geographic information systems (GIS) provide another alternative for interpreting sediment quality data. Using this approach, the matching sediment chemistry, sediment toxicity, and benthic invertebrate structure data are georeferenced in a relational database. Subsequent overlay mapping of the information on the three or more types of indicators facilitates identification of the areas that have various degrees of concordance among the indicators. In this way, it is possible to rank the relative priority of the various reaches in the study area. For example, the reaches in which the majority of sediment samples exhibit elevated chemistry, significant toxicity, and degraded benthos would be considered the highest priority for developing and implementing sediment restoration options. In contrast, those reaches in which a high proportion of samples are relatively uncontaminated, non-toxic, and have normal benthos would be the highest priority for ongoing protection. Other management actions (e.g., further investigation) may be needed in the reaches with one or two indicators showing that the sediments have been degraded. This type of ranking approach can also be applied to non-matching data that have been collected over a number of years.

7.1.2 Integration of Information on Multiple Indicators of Sediment Quality in the Assessment of Impacts on Wildlife

In addition to effects on sediment-dwelling organisms, contaminated sediments have the potential to adversely affect a variety of aquatic-dependent wildlife species, including fish, amphibians, reptiles, birds, and mammals. MacDonald and Ingersoll (2000) evaluated a total of five indicators for determining the potential effects of contaminated sediments on wildlife, including sediment toxicity to fish, fish health, fish community status, sediment chemistry, and tissue chemistry. For most assessments of the effects of contaminated sediments on wildlife species, measures of sediment chemistry, fish community status, and tissue chemistry are the primary indicators evaluated, as sediment toxicity tests with fish and fish health assessments are not routinely reported in assessments of sediment quality conditions. Effects on other wildlife species, such as amphibians, reptiles, birds, and mammals, are often evaluated relative to either sediment chemistry (i.e., by applying bioaccumulation-based SQGs) or fish tissue chemistry (i.e., by applying TRGs for consumption by piscivorus wildlife). The biggest challenge relative to the evaluation of effects of contaminated sediments on fish populations is the mobility of fish within the assessment area. As such, it is difficult to directly link elevated concentrations of contaminants in sediment to effects on fish. Nevertheless, general patterns of sediment contamination within a groups of stations and fish populations samples from the same geographic area can be used to link contaminated sediments to adverse affects on fish (Chapter 6 of Volume III).

The continency table presented in Table 7.4 presents the eight possible outcomes for interpreting the correspondence among measures of sediment chemistry, fish community status, and tissue chemistry relative to the potential for impacts of contaminated sediment on wildlife. Note that if laboratory toxicity tests with fish were conducted with sediments from a station, a contingency table similar to Table 7.1 could be used to evaluate relationships between sediment toxicity, sediment chemistry, and fish community status. Similarly, if fish health was evaluated, a contingency table similar to Table 7.4 could be used (i.e., substitution of fish community status with fish health).

If each of the three indicators listed in Table 7.4 are positive (i.e., bioaccumulation-based SQGs are exceeded, fish community status is impaired, and TRGs are exceeded; line 1 in Table 7.4), it is likely that wildlife are being impacted as a result of sediment contamination in the portion of the assessment area being evaluated. Alternatively, if all three of these indicators are not positive (line 2 in Table 7.4), it is unlikely that wildlife in the assessment area have not been impacted (assuming that these three indicators are representative surrogates for all wildlife inhabiting the portion of the assessment area being evaluated). Again, these comparisons of fish community status (or fish health) and tissue chemistry are often made across multiple stations sampled for sediment chemistry, to account for the fact that fish migrate among stations. Impacts may be identified on fish community status and/or tissue chemistry without an indication of elevated sediment chemistry (lines 4, 5, and 7 in Table 7.4). In these instances, effects on wildlife are probably not due to sediment contamination within the stations being evaluated (tissue residues maybe due to exposure from other sites or other media types). Alternatively, there may be elevated chemistry without noticeable impacts on fish community status or tissue chemistry (line 3 in Table 7.4). In this instance, it may be that fish are not in direct contact with the sediments or the sediment-dwelling organisms from the stations being sampled. Finally, impacts may be identified with sediment chemistry and either fish community status or tissue chemistry (lines 6 and 8 in Table 7.4). In these instances, impacts on sediment quality on wildlife are likely resulting either through direct toxic effects (line 6) or through exceedances of TRGs for piscivorus wildlife (line 8).

7.1.3 Integration of Information on Multiple Indicators of Sediment Quality in the Assessment of Impacts on Human Health

Humans may be exposed to sediment-associated contaminants via several routes of exposure including direct contact with sediment (i.e., wading), through ingestion of surface water contaminated by sediments, or through consumption of shellfish, fish, and/or other wildlife species exposed to contaminated sediments (Chapter 6 of Volume III). Crane (1996) described procedures for evaluating potential human health effects associated with direct contact with contaminated sediment, through ingestion of water contaminated by sediment, and through the consumption of contaminated fish. The contingency table in Table 7.5 addresses the assessment of potential dietary impacts on human health associated with contaminated sediments, as evaluated based on exceedances of sediment chemistry targets (bioaccumulation-based SQGs for human health) or exceedances of tissue chemistry targets (TRGs or fish consumption advisories for human health).

In instances where sediment chemistry and tissue chemistry are elevated in the assessment area (line 1 in Table 7.5), it is likely that sediment contamination has the potential to impact human health. Additionally, when sediment chemistry is elevated in the assessment area above bioaccumulation-based SQGs for humans but tissue chemistry targets are not exceeded (line 3 in Table 7.5), it is possible that there are impacts on human health. In this instance, there may be wildlife in the assessment area exposed to the contaminated sediment that were not sampled for tissue chemistry. Tissue chemistry may be elevated without substantial elevation in sediment chemistry (line 4 in Table 7.5). In this instance, impacts on human health are possible, but organisms may not be exposed to sediments from the sampling stations.

7.2 Summary

Contaminated sediments have the potential to adversely affect sediment -dwelling organisms, wildlife, or human health. Whenever practical, multiple lines of evidence (i.e., data on multiple indicators of sediment quality conditions) should be used to assess the quality of freshwater sediments. Procedures for determining if individual lines of evidence indicate that the beneficial uses of freshwater sediments are being impaired are described in Chapters 2 to 6 of Volume III. The contingency tables presented in this chapter provide a basis for integrating the information on multiple indicators of sediment quality conditions and, in so doing, supporting informed decisions regarding the management of contaminated sediments.

Chapter 8. References

Adams, W.J. 1995. Assessment of the significance of contaminants present in the sediments of the West Branch of the Grand Calumet River. Prepared for the Department of Justice. Prepared by ABC Laboratories, Inc. Columbia, Missouri.

Alexander, C.R., R.G. Smith, F.D. Calder, S.J. Schropp, and H.L. Windom. 1993. The historical record of trace metal enrichment in two Florida estuaries. Estuaries 16:627-636.

Ankley, G.T., P.M. Cook, A.R. Carlson, D.J. Call, J.A. Swenson, H.F. Corcoran, and R.A. Hoke. 1992. Bioaccumulation of PCBs from sediments by oligochaetes and fishes: Comparison of laboratory and field studies. Canadian Journal of Fisheries and Aquatic Sciences 49:2080-2085.

Ankley, G.T., D.M. Di Toro, D.J. Hansen, and W.J. Berry. 1996. Technical basis and proposal for deriving sediment quality criteria for metals. Environmental Toxicology and Chemistry 15:2056-2066.

Antcliffe, B.L., D. Keiser, G. Lawrence, W.L. Lockhart, D.A. Metner, and J.A.J. Thompson. 1997. Monitoring of mountain whitefish (*Prosopium williamsoni*) from the Columbia River system near Castlegar, British Columbia: Final assessment of fish health and contaminants, July 1996. Canadian Technical Report of Fisheries and Aquatic Sciences 2184. Fisheries and Oceans Canada. Vancouver, British Columbia.

ASTM (American Society for Testing and Materials). 2001a. Standard test methods for measuring the toxicity of sediment-associated contaminants with freshwater invertebrates. E1706-00. *In:* ASTM 2001 Annual Book of Standards Volume 11.05. West Conshohocken, Pennsylvania.

ASTM (American Society for Testing and Materials). 2001b. Standard guide for designing biological tests with sediments. E1525-94a. *In:* ASTM 2001 Annual Book of Standards Volume 11.05. West Conshohocken, Pennsylvania.

ASTM (American Society for Testing and Materials). 2001c. Standard guide for collection, storage, characterization, and manipulation of sediments for toxicological testing. E1391-94. *In:* ASTM 2001 Annual Book of Standards Volume 11.05. West Conshohocken, Pennsylvania.

ASTM (American Society for Testing and Materials). 2001d. Standard guide for determination of the bioaccumulation of sediment-associated contaminants by benthic invertebrates. E1688-00a. *In:* ASTM 2001 Annual Book of Standards Volume 11.05. West Conshohocken, Pennsylvania.

ASTM (American Society for Testing and Materials). 2001e. Standard guide for conducting early life-stage toxicity tests with fishes. E1241-98a. *In:* ASTM 2001 Annual Book of Standards Volume 11.05. West Conshohocken, Pennsylvania.

ASTM (American Society for Testing and Materials). 2001f. Standard guide for conducting acute toxicity tests on test materials with fishes, macroinvertebrates, and amphibians. E729-96. *In:* ASTM 2001 Annual Book of Standards Volume 11.05. West Conshohocken, Pennsylvania.

ASTM (American Society for Testing and Materials). 2001g. Standard guide for conducting *Daphnia magna* life-cycle toxicity tests. E1193-97. *In:* ASTM 2001 Annual Book of Standards Volume 11.05. West Conshohocken, Pennsylvania.

ASTM (American Society for Testing and Materials). 2001h. Standard guide for conducting three-brood, renewal toxicity tests with *Ceriodaphnia dubia*. E1295-89. *In:* ASTM 2001 Annual Book of Standards Volume 11.05. West Conshohocken, Pennsylvania.

Barrick, R., S. Becker, R. Pastorok, L. Brown, and H. Beller. 1988. Sediment quality values refinement: 1988 update and evaluation of Puget Sound AET. Prepared by PTI Environmental Services for Environmental Protection Agency. Bellevue, Washington.

Bayne, B.L., D.A. Brown, K. Burns, D.R. Dixon, A. Ivanovici, D.R. Livingstone, D.M. Lowe, M.N. Moore, A.R.D. Stebbing, and J. Widdows. 1985. The effects of stress and pollution on marine mammals. Praeger Publishers. New York, New York. (As cited in Schmitt *et al.* 2000).

Beltman, D. and J. Lipton. 1998. Fish consumption advisories in the Lower Fox River/Green Bay assessment area. Prepared for United States Fish and Wildlife Service, United States Department of the Interior, and United States Department of Justice. Prepared by Stratus Consulting. Boulder, Colorado.

Berry, W.J., D.J. Hansen, J.D. Mahoney, D.L. Robson, D.M. DiToro, D.P. Shipley, B. Rogers, J.M. Corbin, and W.S. Boothman. 1996. Predicting the toxicity of metal-spiked laboratory sediments using acid-volatile sulfide and interstitial water normalizations. Environmental Toxicology and Chemistry 15:2067-2079.

Bolton, S.H., R.J. Breteler, B.W. Vigon, J.A. Scanlon, and S.L. Clark, 1985. National perspective on sediment quality. Prepared for the United States Environmental Protection Agency. Washington, District of Columbia. 194 pp.

Brunson, E.L., T.J. Canfield, F.J. Dwyer, N.E. Kemble, and C.G. Ingersoll. 1998. Assessing bioaccumulation of contaminants from sediments from the upper Mississippi River using field-collected oligochaetes and laboratory-exposed *Lumbriculus variegatus*. Archives of Environmental Contamination and Toxicology 35:191-201.

Burton, G.A. and B.L. Stemmer. 1988. Evaluation of surrogate tests in toxicant impact assessments. Toxicity Assessment: An International Journal 3:255-269.

Burton, G.A., C.G. Ingersoll, L.A.C. Durnett, M. Henry, M.L. Hinman, S.J. Klaine, P.F. Landrum, P. Ross, and M. Tuchman. 1996. A comparison of sediment toxicity test methods at three Great Lakes areas of concern. Journal of Great Lakes Research 22(3):495-511.

Canfield, T.J., N.E. Kemble, W.G. Brumbaugh, F.J. Dwyer, C.G. Ingersoll, and J.F. Fairchild. 1994. Use of benthic invertebrate community structure and the sediment quality triad to evaluate metal-contaminated sediment in the upper Clark Fork River, MT. Environmental Toxicology and Chemistry 13:1999-2012.

Canfield, T.J., F.J. Dwyer, J.F. Fairchild, P.S. Haverland, C.G. Ingersoll, N.E. Kemble, D.R. Mount, T.W. La Point, G.A. Burton, and M.C. Swift. 1996. Assessing contamination in Great Lakes sediments using benthic invertebrate communities and the sediment quality triad approach. Journal of Great Lakes Research 22:565-583.

Canfield, T.J., E.L. Brunson, F.J. Dwyer, C.G. Ingersoll, and N.E. Kemble. 1998. Assessing sediments from the upper Mississippi River navigational pools using benthic invertebrates and the sediment quality triad. Archives of Environmental Contamination and Toxicology 35:202-212.

Carr, R.S., D.C. Chapman, B.J. Presley, J.M. Biedenbach, L. Robertson, P. Boothe, R. Kilada, T. Wade, P. Montagna. 1996. Sediment pore water toxicity assessment studies in the vicinity of offshore oil and gas production platforms in the Gulf of Mexico. Canadian Journal of Fisheries and Aquatic Sciences 53:2618-2628.

Carr, R.S., P.A. Montagna, J.M. Biedenbach, R. Kalke, M.C. Kennicutt, R. Hooten and G. Cripe. 2000. Impact of Storm-water outfalls on sediment quality in Corpus Christi Bay, Texas, USA. Environmental Toxicology and Chemistry 19(3):561-574.

Carvalho, A. and S.J. Schropp. 2001. Development of an interpretive tool for assessment of metal enrichment in Florida freshwater sediment. Taylor Engineering Ltd. Jacksonville, Florida. 38 pp.

CCME (Canadian Council of Ministers of the Environment). 1999. Canadian environmental quality guidelines. Guidelines and Standards Division. Environment Canada. Winnipeg, Manitoba.

CCREM (Canadian Council of Resource and Environment Ministers). 1987. Canadian water quality guidelines. Prepared by the CCME Task Force on Water Quality Guidelines. Ottawa, Canada.

Chapman, P.M. 1992. Sediment quality triad approach. *In:* Sediment Classification Methods Compendium. EPA 823-R-92-006. Office of Water. United States Environmental Protection Agency. Washington, District of Columbia.

Chapman, P.M., M. Cano, A.T. Fritz, C. Gaudet, C.A. Menzie, M. Sprenger, and W.A. Stubblefield. 1997. Workgroup summary report on contaminated site cleanup decisions. *In*: Ecological Risk Assessments of Contaminated Sediment. SETAC Press. Pensacola, Florida. 389 pp.

Chappie, D.J. and G.A. Burton, Jr. 1997. Optimization of *in situ* bioassays with *Hyalella azteca* and *Chironomus tentans*. Environmental Toxicology and Chemistry 16:559-564.

Clements, W.H., D.S. Cherry, and J.H. Van Hassel. 1992. Assessment of the impact of heavy metals on benthic communities at the Clinch River (Virginia): Evaluation of an index of community sensitivity. Canadian Journal of Fisheries and Aquatic Science 49:1686-1694.

Cook, P.M., A.R. Batterman, and K.R. Lodge. 1989. Laboratory measurement of lake trout bioaccumulation of 2,3,7,8-TCDD from Lake Ontario sediment. 32nd Conference on Great Lakes Research at the University of Wisconsin, Madison, Wisconsin. United States Environmental Protection Agency. Environmental Research Laboratory. Duluth, Minnesota.

Cook, P.M., A.R. Carlson, and H. Lee. 1992. Tissue residue approach. *In:* Sediment Classification Methods Compendium. EPA 823-R-92-006. Office of Water. United States Environmental Protection Agency. Washington, District of Columbia.

Crane, J.L. 1996. Carcinogenic human health risks associated with consuming contaminated fish from five Great Lakes Areas of Concern. Journal of Great Lakes Research 22:653-668.

Crane, J.L., D.D. MacDonald, C.G. Ingersoll, D.E. Smorong, R.A. Lindskoog, C.G. Severn, T.A. Berger, and L.J. Field. 2000. Development of a framework for evaluating numerical sediment quality targets and sediment contamination in the St. Louis River Area of Concern. EPA 905-R-00-008. Great Lakes National Program Office. United States Environmental Protection Agency. Chicago, Illinois. 107 pp. + appendices.

Crane, M. and L. Maltby. 1991. The lethal and sub-lethal responses of *Gammarus pulex* to stress: Sensitivity and variation in an *in situ* bioassay. Environmental Toxicology and Chemistry 10:1331-1339.

Crane, M., P. Delaney, P. Parker, C. Walker, and S. Watson. 1995a. The effect of Malathion 60 on *Gammarus pulex* (L.) below watercress beds. Environmental Toxicology and Chemistry 14:1181-1188.

Crane, M., P. Delaney, C. Mainstone, and S. Clarke. 1995b. Measurement by *in situ* bioassay of water quality in an agricultural catchment. Water Research 29:2441-2448.

Crane, M., I. Johnson, and L. Maltby. 1996. *In situ* assays for monitoring the toxic impacts of waste in rivers. *In:* Toxic Impacts of Wastes on the Aquatic Environment. J.F. Tapp, S.M. Hunt, and J.R. Wharfe (Eds.). The Royal Society of Chemistry. Cambridge, United Kingdom. pp. 116-124.

Crane, M., C. Attwood, D. Sheahan, and S. Morris. 1999. Toxicity and bioavailability of the organophosphorus pesticide pirimiphos methyl to the freshwater amphipod *Gammarus pulex* L. in laboratory and mesocosm systems. Environmental Toxicology and Chemistry 18:1456-1461.

Crane, M., M. Higman, K. Olsen, P. Simpson, A. Callaghan, T. Fisher, and R. Kheir. 2000. An *in situ* system for exposing aquatic invertebrates to contaminated sediments. Environmental Toxicology and Chemistry 19:2715-2719.

Cubbage, J., D. Batts, and S. Briedenbach. 1997. Creation and analysis of freshwater sediment quality values in Washington State. Environmental Investigations and Laboratory Services Program. Washington Department of Ecology. Olympia, Washington.

Davis, W.S. and J.E. Lathrop. 1992. Freshwater benthic macroinvertebrate community structure and function. *In:* Sediment Classification Methods Compendium. EPA 823-R-92-006. Office of Water. United States Environmental Protection Agency. Washington, District of Columbia.

Di Toro, D.M. and J.A. McGrath. 2000. Technical basis for narcotic chemicals and polycyclic aromatic hydrocarbon criteria. II. Mixtures and sediments. Environmental Toxicology and Chemistry 19(8):1971-1982.

Di Toro, D.M., C.S. Zarba, D.J. Hansen, W.J. Berry, R.C. Swartz, C.E. Cowan, S.P. Pavlou, H.E. Allen, N.A. Thomas, and P.R. Paquin. 1991. Technical basis for establishing sediment quality criteria for non-ionic organic chemicals using equilibrium partitioning. Environmental Toxicology and Chemistry 10(12):1541-1583.

Di Toro, J.D. Mahony, D.J. Hansen, K.J. Scott, A.R. Carlson, and G.T. Ankley. 1992. Acid volatile sulfide predicts the acute toxicity of cadmium and nickel in sediments. Environmental Science and Technology 26(1):96-101.

Dunn, G.W. 1989. An approach used to establish site-specific water quality indicators on interprovincial streams. Proceedings of the CCME Workshop on the Development of Water Quality Objectives. Canadian Council of Ministers of the Environment. Ottawa, Canada.

EC and HC (Environment Canada and Health Canada). 1994. Nickel and its compounds. Priority substances list assessment report. Canadian Environmental Protection Act. Ottawa, Canada.

EC and MENVIQ (Environment Canada and Ministere de l'Envionnement du Quebec). 1992. Interim criteria for quality assessment of St. Lawrence River sediment. ISBN 0-662-19849-2. Environment Canada. Ottawa, Ontario.

Endicott, D., W. Richardson and D.M. Di Toro. 1989. Lake Ontario TCDD modelling report. Large Lakes Research Station. Environmental Research Laboratory-Duluth. United States Environmental Protection Agency. Grosse Ile, Minnesota. 94 pp.

Environment Canada. 1997a. Biological test method: Test for growth and survival in sediment using larvae of freshwater midges (*Chironomus tentans* or *Chironomus riparius*). Environment Canada, Ottawa, Ontario, Technical report EPS 1/RM/32.

Environment Canada. 1997b. Biological test method: Test for growth and survival in sediment using the freshwater amphipod *Hyalella azteca*. Environment Canada, Ottawa, Ontario. Technical report EPS 1/RM/33.

Field, L.J., D.D. MacDonald, S.B. Norton, C.G. Severn, and C.G. Ingersoll. 1999. Evaluating sediment chemistry and toxicity data using logistic regression modeling. Environmental Toxicology and Chemistry 18(6):1311-1322.

Field, L.J., D.D. MacDonald, S.B. Norton, C.G. Ingersoll, C.G. Severn, D.E. Smorong, and R.A. Lindskoog. 2002. Predicting amphipod toxicity from sediment chemistry using logistic regression models. Submitted to Environmental Toxicology and Chemistry. In press.

Gardner W.S., W.A. Frez, E.A. Cichocki, and C.C. Parrish. 1985. Micromethods for lipids in aquatic invertebrates. Limnology and Oceanography 30:1099-1105.

Goyette, D., D. Brand, and M. Thomas. 1988. Prevalence of idiopathic liver lesions in English sole and epidermal abnormalities in flatfish from Vancouver Harbour, British Columbia, 1986. Regional Program Report 87-09. Environment Canada. Vancouver, British Columbia. 48 pp.

Grady, A.W., R.M. McLaughlin, C.W. Caldwell, C.J. Schmitt, and D.J. Stalling. 1992. Flow cytometry, morphometry, and histopathology as biomarkers of benzo(a)pyrene exposure in brown bullheads (*Ameiurus nebulosos*). Journal of Applied Toxicology 12:165-177. (As cited in Schmitt *et al.* 2000).

Griffith, G.E., J.M. Omernik, C.M. Rohm, and S.M. Pierson. 1994. Florida Regionalization Project. Environmental Research Laboratory. Office of Research and Development. United States Environmental Protection Agency. Corvallis, Oregon.

Hansen, D.J., W.J. Berry, J.D. Mahoney, W.S. Boothman, D.M. DiToro, D.L. Robson, G.T. Ankley, D. Ma, Q. Yan, and C.E. Pesch. 1996. Predicting the toxicity of metal-contaminated field sediments using interstitial concentration of metals and acid-volatile sulfide normalizations. Environmental Toxicology and Chemistry 15:2080-2094.

Harkey, G.A., P.F. Landrum, and S.J. Klaine. 1994. Comparison of whole-sediment, elutriate and pore water exposures for use in assessing sediment-associated organic contaminants in bioassays. Environmental Toxicology and Chemistry 13:1315-1329.

Hayward, J.M. 2002. Field validation of laboratory sediment toxicity tests with *Hyalella azteca* and *Chironomus tentans*. PhD Thesis. University of Missouri. Columbia, Missouri.

Hilsenhoff, WL. 1982. Using a biotic index to evaluate water quality in streams. Technical Bulletin No. 132. Department of Natural Resources. Madison, Wisconsin.

Hilsenhoff, WL. 1987. An improved biotic index of organic stream pollution. Great Lakes Entomologist 20:31-39.

IDEM (Indiana Department of Environmental Management). 1992. Indiana 305(b) report 1990-91. Office of Water Management. Indianapolis, Indiana. 227 pp.

IJC (International Joint Commission). 1988. Procedures for the assessment of contaminated sediment problems in the Great Lakes. Prepared by the Sediment Subcommittee and its Assessment Work Group. Great Lakes Regional Office. Windsor, Ontario. 140 pp.

Ingersoll, C.G. and D.D. MacDonald. 1999. An assessment of sediment injury in the West Branch of the Grand Calumet River. Volume I. Prepared for Environmental Enforcement Section. Environment and Natural Resources Division. United States Department of Justice. Washington, District of Columbia. 161 pp.

Ingersoll, C.G., T. Dillon, and R.G. Biddinger (Eds.). 1997. Methodological uncertainty in sediment ecological risk assessment. *In*: Ecological Risk Assessments of Contaminated Sediment. SETAC Press. Pensacola, Florida. 389 pp.

Ingersoll, C.G., D.D. MacDonald, N. Wang, J.L. Crane, L.J. Field, P.S. Haverland, N.D. Kemble, R.A. Lindskoog, C.G. Severn, and D.E. Smorong. 2001a. Predictions of sediment toxicity using consensus-based freshwater sediment quality guidelines. Archives of Environmental Contamination and Toxicology 41:8-21.

Ingersoll, C.G., D.D. MacDonald, W.G. Brumbaugh, S.D. Ireland, B.T. Johnson, N.E. Kemble, J.L. Kunz, and T.W. May. 2001b. Toxicity assessment of sediments from the Grand Calumet River and Indiana Harbor Canal in northwestern Indiana. Archives of Environmental Contamination and Toxicology. In press.

Ingersoll, C.G., E.L. Brunson, N. Wang, F.J. Dwyer, J. Huckins, J. Petty, G.T. Ankley, D.R. Mount, and P.F. Landrum. 2001c. Uptake and depuration of sediment-associated contaminants by the oligochaete, *Lumbriculus variegatus*. Manuscript in review.

Ireland, D.S., G.A. Burton, Jr., and G.G. Hess. 1996. *In situ* toxicity evaluations of turbidity and photoinduction of polycyclic aromatic hydrocarbons. Environmental Toxicology and Chemistry 15:574-581.

Jarvinen, A.W. and G.T. Ankley. 1999. Linkage of effects to tissue residues: Development of a comprehensive database for aquatic organisms to inorganic and organic chemicals. SETAC Technical Publication Series. Society of Environmental Toxicology and Chemistry. Pensacola, Florida. 364 pp.

Johnson B.T. and E.R. Long. 1998. Rapid toxicity assessment of sediments from large estuarine ecosystems: A new tandem *in vitro* testing approach. Environmental Toxicology and Chemistry 17:1099-1106.

Kadeg, R.D., S.P. Pavlou, and A.S. Duxbury. 1986. Sediment criteria methodology validation: Elaboration of sediment normalization theory for non-polar hydrophobic organic chemicals - Final Report. Prepared for Office of Water Regulations and Standards. United States Environmental Protection Agency. Washington, District of Columbia.

Karr, J.R. and E.W. Chu. 1997. Biological monitoring and assessment: Using multimetric indexes effectively EPA 235-R97-001. University of Washington. Seattle, Washington.

Karr, J.R. and E.W. Chu. 1999. Restoring life to running waters: Better biological monitoring. Island Press. Washington, District of Columbia. 206 pp.

Kemble, N.E., F.J. Dwyer, J.L. Kunz, and C.G. Ingersoll. In preparation. Relative sensitivity of endpoints measured in long-term water-only exposures with the amphipod, *Hyalella azteca* and the midge, *Chironomus tentans*. To be submitted to Environmental Toxicology and Chemistry.

Kemble, N.E., W.G. Brumbaugh, E.L. Brunson, F.J. Dwyer, C.G. Ingersoll, D.P. Monda, and D.F. Woodward. 1994. Toxicity of metal-contaminated sediments from the upper Clark Fork River, MT to aquatic invertebrates in laboratory exposures. Environmental Toxicology and Chemistry 13:1985-1997.

Krantzberg, G., M.A. Zarull, and J.H. Hartig. 2001. Sediment management: Ecological and ecotoxicological effects must direct actions. Water Quality Research Journal of Canada 36(3):367-376.

Landrum, K.E. 1995. Trace metal variability of estuarine sediments, St. Bernard Geomorphic Region, Louisiana. Gulf Coast Association of Geological Societies Transactions: XLV:365-370. Tulane University, Department of Geology. New Orleans, Louisiana.

Long, E.R. and P. Chapman. 1985. A sediment quality triad: Measurements of sediment contamination, toxicity, and infaunal community composition in Puget Sound Marine Pollution Bulletin 16:405-415.

Long, E.R. and L.G. Morgan. 1991. The potential for biological effects of sediment-sorbed contaminants tested in the National Status and Trends Program. NOAA Technical Memorandum NOS OMA 52. National Oceanic and Atmospheric Administration. Seattle, Washington. pp. 175 + appendices.

Long, E.R. and C.J. Wilson. 1997. On the identification of toxic hot spots using measures of the sediment quality triad. Marine Pollution Bulletin 34:373-374.

Long, E.R. and D.D. MacDonald. 1998. Recommended uses of empirically derived, sediment quality guidelines for marine and estuarine ecosystems. Human and Ecological Risk Assessment 4:1019-1039.

Long, E.R., D.D. MacDonald, S.L. Smith, and F.D. Calder. 1995. Incidence of adverse biological effects within ranges of chemical concentrations in marine and estuarine sediments. Environmental Management 19:81-97.

Long, E.R., D.D. MacDonald, J.C. Cubbage, and C.G. Ingersoll. 1998a. Predicting the toxicity of sediment-associated trace metals with SEM-AVS concentrations and dry weight-normalized concentrations - A critical comparison. Environmental Toxicology and Chemistry 17:972-972.

Long, E.R., L.J. Field, and D.D. MacDonald. 1998b. Predicting toxicity in marine sediments with numerical sediment quality guidelines. Environmental Toxicology and Chemistry 17:714-727.

Loring, D.H. 1991. Normalization of heavy-metal data from estuarine and coastal sediments. ICES Journal of Marine Science 48:101-115.

Luoma, S.N. and N. Fisher. 1997. Uncertainties in assessing contaminant exposure from sediments. *In:* Ecological Risk Assessment of Contaminated Sediment. C.G. Ingersoll, T. Dillon, and G.R. Biddinger (Eds.). SETAC Press. Pensacola, Florida. pp. 211-237.

MacDonald, D.D. 1993. Canadian environmental quality guidelines for polychlorinated dibenzo-*p*-dioxins and polychlorinated dibenzo furans. Report prepared for EcoHealth Branch, Environment Canada. Ottawa Canada. 212 pp.

MacDonald, D.D. 1994a. Approach to the assessment of sediment quality in Florida coastal waters. Volume 1: Development and evaluation of sediment quality assessment guidelines. Report prepared for Florida Department of Environmental Protection. Tallahassee, Florida.

MacDonald, D.D. 1994b. Approach to the assessment of sediment quality in Florida coastal waters. Volume 2: Applications of the sediment quality assessment guidelines. Report prepared for Florida Department of Environmental Protection. Tallahassee, Florida.

MacDonald, D.D. 1994c. Canadian sediment quality guidelines for polycyclic aromatic hydrocarbons. Prepared for Evaluation and Interpretation Branch. Environment Canada. Ottawa, Canada. 195 pp.

MacDonald, D.D. 1995. Canadian sediment quality guidelines for polychlorinated biphenyls: Draft. Prepared for the Guidelines Division. Environment Canada. Ottawa, Canada.

MacDonald, D.D. 1997. Tampa Bay sediment quality workshop: Setting targets and defining management strategies. Final summary report. Technical Publication #06-97. Prepared by MacDonald Environmental Sciences Ltd. Prepared for the Tampa Bay National Estuary Program. St. Petersburg, Florida.

MacDonald, D.D. 1998. An ecosystem-based framework for assessing sediment quality in the Great Lakes Basin. A short course on collection, analysis, and interpretation of sediment quality data: Applications of sediment quality guidelines (SQGs) and various companion tools. Great Lakes National Program Office. United States Environmental Protection Agency. Chicago, Illinois.

MacDonald, D.D. and A. Sobolewski. 1993. Recommended procedures for developing site-specific environmental quality remediation objectives for contaminated sites. Report prepared for EcoHealth Branch, Environment Canada. Ottawa Canada. 85 pp.

MacDonald, D.D. and M. Macfarlane. 1999. Criteria for managing contaminated sediment in British Columbia. Prepared pursuant to Section 26(a) of the Waste Management Act. Prepared for Ministry of Environment, Lands, and Parks. Victoria, British Columbia.

MacDonald, D.D. and C.G. Ingersoll. 2000. An assessment of sediment injury in the grand Calumet River, Indiana Harbor Canal, Indiana Harbor, and the nearshore areas of Lake Michigan. Volume I. Prepared for the U.S. Fish and Wildlife Service. Bloomington, Indiana. 238 pp.

MacDonald, D.D., R.S. Carr, F.D. Calder, E.R. Long, and C.G. Ingersoll. 1996. Development and evaluation of sediment quality guidelines for Florida coastal waters. Ecotoxicology 5:253-278.

MacDonald, D.D., L.M. DiPinto, J. Field, C.G. Ingersoll, E.R. Long, and R.C. Swartz. 2000a. Development and evaluation of consensus-based sediment effect concentrations for polychlorinated biphenyls (PCBs). Environmental Toxicology and Chemistry 19:1403-1413.

MacDonald, D.D., C.G. Ingersoll, and T.A. Berger. 2000b. Development and evaluation of consensus-based sediment quality guidelines for freshwater ecosystems. Archives of Environmental Contamination and Toxicology 39:20-31.

MacDonald, D.D., D.R. Moore, A. Pawlisz, D.E. Smorong, R.L. Breton, D.B. MacDonald, R. Thompson, R.A. Lindskoog and M.A. Hanacek. 2001a. Calcasieu estuary remedial investigation/feasibility study (RI/FS): Baseline ecological risk assessment (BERA). Volume I. Baseline problem formulation (Draft). Prepared for: CDM Federal Programs Corporation. Dallas, Texas.

MacDonald, D.D., D.R. Moore, A. Pawlisz, D.E. Smorong, R.L. Breton, D.B. MacDonald, R. Thompson, R.A. Lindskoog and M.A. Hanacek. 2001b. Calcasieu estuary remedial investigation/feasibility study (RI/FS): Baseline ecological risk assessment (BERA). Volume 2. Baseline problem formulation: Appendices (Draft). Prepared for: CDM Federal Programs Corporation. Dallas, Texas.

Malins, D.C., M.M. Krahn, M.S. Myers, L.D. Rhodes, D.W. Brown, C.A. Krone, B.B. McCain, and S.-L. Chan. 1985. Toxic chemicals in sediments and biota from a creosote-polluted harbor: Relationships with hepatic neoplasms and other hepatic lesions in English sole (*Parophrys vetulus*). Carcinogenesis 6(10):1463-1469.

Maltby, L. and M. Crane. 1994. Responses of *Gammarus pulex* (Amphipoda, Crustacea) to metalliferous effluents: Identification of toxic components and the importance of interpopulation variation. Environmental Pollution 84:45-52.

McCarthy, J.F. and L.R. Shugart (Eds.). 1990. Biomarkers of Environmental Contamination. Lewis Publishers. Boca Raton, Florida.

Menzie, C., M.H. Henning, J. Cura, K. Finkelstein, J. Gentile, J. Maughan, D. Mitchell, S. Petron, B. Potocki, S. Svirsky, and P. Tyler. 1996. Special report of the Massachusetts weight-of-evidence workgroup: A weight-of-evidence approach for evaluating ecological risks. Human Ecological Risk Assessment 2:277-304.

Menzie, C.A. 1997. Perspectives on sediment ecological risk analysis for hazardous waste sites. *In*: C.G. Ingersoll, T. Dillon, and G.R. Biddinger (Ed.). Ecological Risk Assessment of Contaminated Sediments. SETAC Press. Pensacola, Florida. .

Mudroch, A. and J.M. Azcue. 1995. Manual of Aquatic Sediment Sampling. Lewis Publishers. Boca Raton, Florida.

Neff, J.M., D.J. Bean, B.W. Cornaby, R.M. Vaga, T.C. Gulbransen, and J.A. Scanlon. 1986. Sediment quality criteria methodology validation: Calculation of screening level concentrations from field data. Prepared for Environmental Protection Agency Region V. Washington, District of Columbia. 225 pp.

Nener, J., D. Keiser, J.A.J. Thompson, W.L. Lockhart, D.A. Metner, and R. Roome. 1995. Monitoring of mountain whitefish (*Prosopium williamsoni*) from the Columbia River system near Castlegar, British Columbia: Health parameters and contaminants in 1992. Canadian Manuscript Report of Fisheries and Aquatic Sciences 2036. Fisheries and Oceans Canada. Vancouver, British Columbia.

Newell, A.J., D.W. Johnson, and L.K. Allen. 1987. Niagara River biota contamination project: Fish flesh criteria for piscivorus wildlife. Technical Report 87-3. New York State Department of Environmental Conservation. Albany, New York. 182 pp.

NYSDEC (New York State Department of Environmental Conservation). 1999. Technical guidance for screening contaminated sediments. Division of Fish, Wildlife and Marine Resources. Albany, New York. 39 pp.

OEPA (Ohio Environmental Protection Agency). 1988a. Biological criteria for the protection of aquatic life: Volume I: The role of biological data in water quality assessment. Ecological Assessment Section. Division of Water Quality Planning and Assessment. Columbus, Ohio.

OEPA (Ohio Environmental Protection Agency). 1988b. Biological criteria for the protection of aquatic life: Volume II: Users manual for biological field assessment of Ohio surface waters. Ecological Assessment Section. Division of Water Quality Planning and Assessment. Columbus, Ohio.

OEPA (Ohio Environmental Protection Agency). 1989. Biological criteria for the protection of aquatic life: Volume III: Standardized biological field sampling and laboratory methods for assessing fish and macroinvertebrate communities. Ecological Assessment Section. Division of Water Quality Planning and Assessment. Columbus, Ohio.

Olsen, A., L. Ellerbeck, T. Fisher, A. Callaghan, and M. Crane. 2001. Variability in acetylcholinesterase and glutathione-s-transferase activities in *Chironomus riparius* Meigen deployed *in situ* at uncontaminated field sites. Environmental Toxicology and Chemistry: In press.

Outridge, P.M., D.D. MacDonald, E. Porter, and I.E. Cuthbert. 1994. An evaluation of the ecological hazards associated with cadmium in the Canadian environment. Environmental Reviews 2:91-107.

Pardue, J.H., R.D. DeLaune, and W.H. Patrick, Jr. 1992. Metal to aluminum correlation in Louisiana coastal wetlands: Identification of elevated metal concentrations. Journal of Environmental Quality 21:539-545.

Pavlou, S.P. 1987. The use of equilibrium partitioning approach in determining safe levels of contaminants in marine sediments. *In:* Fate and effects of sediment-bound chemicals in aquatic systems. K.L. Dickson, A.W. Maki, and W.A. Brungs (Eds.). Proceedings of the Sixth Pellston Workshop. Florissant, Colorado, August 13-17, 1984. Pergamon Press. New York, New York. pp. 388-395.

Pavlou, S.P. and D.P. Weston. 1983. Initial evaluation of alternatives for development of sediment related criteria for toxic contaminants in marine waters (Puget Sound). Phase I: Development of conceptual framework. Final Report. JRB Associates. Bellevue, Washington. 56 pp.

Payne, J.F., J. Kiceniuk, L.L. Fancey, U. Williams, G.L. Fletcher, A. Rahimtula, and B. Fowler. 1988. What is a safe level of polycyclic aromatic hydrocarbons for fish: Subchronic toxicity study of winter flounder (*Pseudopleuronectes americanus*). Canadian Journal of Fisheries and Aquatic Sciences 45:1983-1993.

Persaud, D., R. Jaagumagi, and A. Hayton. 1989. Development of provincial sediment quality guidelines. Water Resources Branch. Ontario Ministry of the Environment. Toronto, Ontario. 19 pp.

Persaud, D., R. Jaagumagi, and A. Hayton. 1993. Guidelines for the protection and management of aquatic sediment quality in Ontario. Standards Development Branch. Ontario Ministry of Environment and Energy. Toronto, Ontario. 27 pp.

Porebski, L. 1999. A short course on collection, analysis, and interpretation of sediment quality data: applications of sediment quality guidelines (SQGs) and various companion tools. September 21 and 22, 1999. State Botanical Garden of Georgia. Athens, Georgia Prepared for the United States Environmental Protection Agency.

Reynoldson T.B. and K.E. Day. 1998. Biological guidelines for the assessment of sediment quality in the Laurentian Great Lakes. NWRI Contribution Number 98-232. National Water Research Institute. Environment Canada. Burlington, Ontario.

Reynoldson T.B. and P. Rodriguez. 1999. Field methods and interpretation for sediment bioassessment. pp. 135-175. *In:* Manual of Bioassessment of Aquatic Sediment Quality. A. Mudroch, J.M. Azcue, and P. Mudroch (Eds.). Lewis Publishers. Boca Raton, Florida.

Reynoldson, T.B., K.E. Day, C. Clark, and D. Milani. 1994. Effect of indigenous animals on chronic end points in freshwater sediment toxicity tests. Environmental Toxicology and Chemistry 13(6):973-977.

Reynoldson T.B., K.E. Day, R.C. Bailey, and R.H. Norris. 1995. Methods for establishing biologically based sediment guidelines for freshwater quality management using benthic assessment of sediment. Australian Journal of Ecology 20:198-219.

Reynoldson T.B., R.H. Norris, V.H. Resh, K.E. Day, and D.M. Rosenberg. 1997. The reference condition: A comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. Journal of the North American Benthological Society 16:833-852.

Reynoldson, T. B., L. Grapentine, M. Zarull, T. Pascoe, L. Richman, C. DeBarros, S. Painter, and J. Anderson. 2000. A support system for decision making using chemical and biological measures. National Water Research Institute. Environment Canada. Burlington, Ontario.

Rosenberg, D.M. and A.P. Wiens. 1976. Effects of sediment addition on macrobenthic invertebrates in a Northern Canadian River. Water Research 12:753-763.

Rosenberg, D.M. and V.H. Resh. 1993. Freshwater biomonitoring and benthic macroinvertebrates. Chapman and Hall. New York. 488 p.

Sample, B.E. and D.M. Opresko. 1996. Toxicological benchmarks for wildlife: 1996 revision. ES/ER/TM-86/R3. Oak Ridge National Laboratory. Oak Ridge, Tennessee.

Sarda, N. and G.A. Burton, Jr. 1995. Ammonia variation in sediments: Spatial, temporal and method-related effects. Environmental Toxicology and Chemistry 14:1499-1506.

Schiewe, M.H., E.G. Hawk, D.I. Actor, and M.M. Krahn. 1985. Use of a bacterial bioluminescence assay to assess toxicity of contaminated sediments. Canadian Journal of Fisheries and Aquatic Science 42:1244-1248.

Schiff, K. and S.B. Weisberg. 1996. Iron as a reference element for determining trace metal enrichment in California Coastal Shelf sediments. SSWRP 1996 Annual Report (www/sccwrp.org/pubs/annrpt/96/ar-07.htm).

Schmitt, C.J., D.E. Tillitt, and V.S. Blazer. 2000. The 1995 and 1997 BEST projects. *In:* Biomonitoring of Environmental Status and Trends (BEST) Program: Selected Methods for Monitoring Chemical Contaminants and their Effects in Aquatic Ecosystems. C.J. Schmitt and G.M. Dethloff (Eds.). Information and Technology Report USGS/BRD/ITR–2000-0005. U.S. Geological Survey. Columbia, Missouri.

Schropp, S.J. and H.L. Windom. 1988. A guide to the interpretation of metal concentrations in estuarine sediments. Coastal Zone Management Section. Florida Department of Environmental Regulation. Tallahassee, Florida. 44 pp. + app.

Schropp, S.J., F.D. Calder, L.C. Burney, and H.L. Windom. 1989. A practical approach for assessing metals contamination in coastal sediments - An example from Tampa Bay. *In:* Proceedings of the Sixth Symposium on Coastal and Ocean Management. July 11 - 14, 1989. Published by the American Society of Civil Engineers. Charleston, North Carolina.

Schropp, S.J., F.G. Lewis, H.L. Windom, J.D. Ryan, F.D. Calder, and L.C. Burney. 1990. Interpretation of metal concentrations in estuarine sediments of Florida using aluminum as a reference element. Estuaries 13:227-235.

Seager, J., I. Milne, and M. Crane. 1991. The application of *in situ* bioassays as ecological indicators for assessing river quality. *In:* Ecological Indicators Vol. 1. D.H. McKenzie, D.E. Hyatt, and V.J. McDonald (Eds.). Elsevier Applied Science. London. pp. 243-258.

Seager, J., I. Milne, G. Rutt, and M. Crane. 1992. Integrated biological methods for river quality assessment. *In:* River Water Quality: Ecological Assessment and Control. P.J. Newman, M.A. Piavaux and R.A. Sweeting (Eds.). CEC. Brussels. pp. 399-415.

SETAC (Society of Environmental Toxicology and Chemistry). 1999. Endocrine disruption in invertebrates: Endocrinology, testing, and assessment. Pensacola, Florida. SETAC Press. 303 p.

Simon, T.P., D. Sparks, S. Newhouse, and R. Dufour. 2000. Aquatic community injury assessment for the Grand Calumet River, Indiana Harbor Canal, and the nearshore areas of Lake Michigan. Natural resource damage assessment. United States Fish and Wildlife Service. Bloomington, Indiana.

Smith, S.L., D.D. MacDonald, K.A. Keenleyside, C.G. Ingersoll, and J. Field. 1996. A preliminary evaluation of sediment quality assessment values for freshwater ecosystems. Journal of Great Lakes Research 22:624-638.

Sobiech, S.A., T.P. Simon, and D.W. Sparks. 1994. Pre-remedial biological and water quality assessment of the East Branch Grand Calumet River, Gary, Indiana. Prepared for United States Environmental Protection Agency Region V. Chicago Illinois. Prepared by United States Fish and Wildlife Service. Bloomington, Indiana. 111 pp.

Spear, P.A. and R.C. Pierce. 1979. Copper in the aquatic environment: Chemistry, distribution and toxicology. NRCC No. 16454. National Research Council of Canada. Ottawa, Canada.

Striplin, B., G. Baun, and G. Bilyard. 1992. Marine benthic community structure assessment. *In:* Sediment Classification Methods Compendium. EPA 823-R-92-006. Office of Water. United States Environmental Protection Agency. Washington, District of Columbia.

Swartz, R.C. 1999. Consensus sediment quality guidelines for PAH mixtures. Environmental Toxicology and Chemistry 18:780-787.

Swartz, R.C., D.W. Schults, T.H. DeWitt, G.R. Ditsworth, and J.O. Lamberson. 1987. Toxicity of fluoranthene in sediment to marine amphipods: A test of the equilibrium partitioning approach to sediment quality criteria. 89th Annual Meeting, Society for Environmental Toxicology and Chemistry. Pensacola, Florida. November 1987. 12 pp.

Swartz, R.C., D.W. Schults, R.J. Ozretich, J.O. Lamberson, F.A. Cole, T.H. DeWitt, M.S. Redmond, and S.P. Ferraro. 1995. PAH: A model to predict the toxicity of polynuclear aromatic hydrocarbon mixtures in field-collected sediment. Environmental Toxicology and Chemistry 14(11):1977-1987.

Tetra Tech, Inc. 1986. Development of sediment quality values for Puget Sound. Volume 1. Puget Sound Dredged Disposal Analysis Report. Seattle, Washington. 129 pp.

Thomann, R.V. 1989. Bioaccumulation model of organic chemical distributions in aquatic food chains. Environmental Science and Technology 23:699-715.

Thursby, G.B., Heltshe, J., and K.J. Scott. 1997. Revised approach to toxicity test acceptability criteria using a statistical performance assessment. Environmental Toxicology and Chemistry 16:1322-1329.

USEPA (United States Environmental Protection Agency). 1989. Assessing human health risks from chemically contaminated fish and shellfish: A guidance manual. EPA-503/8-89-002. Office of Water Regulations and Standards. Washington, District of Columbia. 91 pp + app.

USEPA (United States Environmental Protection Agency). 1990. Sediment investigation Indiana Harbor Canal - Sampling dates: December 6, 7, 10, and 11, 1990. Environmental Services Division. Central District Office. Region V. Chicago, Illinois. 68 pp. + apps.

USEPA (United States Environmental Protection Agency). 1991. Sediment toxicity identification evaluation: Phase I (Characterization), Phase II (Identification) and Phase III (Confirmation). Modifications of effluent procedures. EPA-600/6-91/007. Duluth, Montana.

USEPA (United States Environmental Protection Agency). 1992a. Freshwater benthic macroinvertebrate community structure and function. Sediment classification methods compendium. Chapter 8. EPA 813/R-92/006. Washington, District of Columbia.

USEPA (United States Environmental Protection Agency). 1992b. Marine benthic community structure assessment. Sediment classification methods compendium. Chapter 9. EPA 813/R-92/006. Washington, District of Columbia.

USEPA (United States Environmental Protection Agency). 1993. Water quality criteria summary. Ecological Risk Assessment Branch (WH-586). Health and Ecological Criteria Division. Office of Science and Technology. Washington, District of Columbia. 3 pp.

USEPA (United States Environmental Protection Agency). 1994. Assessment and remediation of contaminated sediments (ARCS) program. Great Lakes National Program Office. EPA 905/B-94/002. Chicago, Illinois.

USEPA (United States Environmental Protection Agency). 1996. Assessment and remediation of contaminated sediments (ARCS) program. Calculation and evaluation of sediment effect concentrations for the amphipod *Hyalella azteca* and the midge *Chironomus riparius*. Great Lakes National Program Office. EPA 905/R-96/008. Chicago, Illinois.

USEPA (United States Environmental Protection Agency). 1997. The incidence and severity of sediment contamination in surface waters of the United States: Volume 1: National Sediment Quality Survey. EPA/823/R-97/006. Washington, DC.

USEPA (United States Environmental Protection Agency). 1999. National recommended water quality criteria - correction. EPA-822-Z-99-001. Office of Water. Washington, District of Columbia.

USEPA (United States Environmental Protection Agency). 2000a. Methods for measuring the toxicity and bioaccumulation of sediment-associated contaminants with freshwater invertebrates. Second edition. EPA 600/R-99/064. Office of Research and Development. Washington, District of Columbia.

USEPA (United States Environmental Protection Agency). 2000b. Prediction of sediment toxicity using consensus-based freshwater sediment quality guidelines. EPA 905/R-00/007. Great Lakes Program Office. Chicago, Illinois.

USEPA (United States Environmental Protection Agency). 2000c. Stressor identification guidance document. EPA-822-B-00-025. Office of Water. Office of Research and Development. Washington, District of Columbia.

USEPA (United States Environmental Protection Agency). 2001. Methods for collection, storage and manipulation of sediments for chemical and toxicological analyses: Technical manual. EPA-823-B-01-002. Office of Water. Washington, District of Columbia.

USEPA and USACE (United States Environmental Protection Agency and United States Army Corps of Engineers). 1998a. Evaluation of dredged material proposed for discharge in water of the US. Testing Manual. EPA 823-B-98-004. United States Environmental Protection Agency. Washington, District of Columbia.

USEPA and USACE (United States Environmental Protection Agency and United States Army Corps of Engineers). 1998b. Method for assessing the chronic toxicity of marine and estuarine sediment-associated contaminants with the amphipod *Leptocheirus plumulosus*. Office of Research and Development. Newport, Oregon (August 1998 draft). (As cited in USEPA 2000b).

USGS (United States Geological Survey). 1998. Contaminant Hazard Reviews. Patuxent Wildlife Research Center. Laurel, Maryland.

USGS (United States Geological Survey). 2000. Biomonitoring and Environmental Status and Trends (BEST) monitoring program: Select methods for monitoring chemical contaminants and their effects in aquatic systems. USGS Information and Technology Report USGS/BRD/ITR 2000/005. Columbia, Missouri. 81 p.

Valette-Silver, N.J. 1993. The use of sediment cores to reconstruct historical trends in contamination of estuarine and coastal sediments. Estuaries 16:577-588.

Veerasingham, M. and M. Crane. 1992. Impact of farm waste on freshwater invertebrate abundance and the feeding rate of *Gammarus pulex* L. Chemosphere 25: 869-874.

Washington Department of Ecology. 1990. Sediment management standards: Chapter 173-204 WAC. Olympia, Washington. 106 pp.

Washington State Department of Health. 1995. Development of sediment quality criteria for the protection of human health. Tier I report. Prepared by Environmental Health Programs. Olympia, Washington.

Washington State Department of Health. 1996. Development of sediment quality criteria for the protection of human health. Tier II report. Olympia, Washington.

Yoder, C.O. and E.T. Rankin. 1995. Biological response signatures and the area of degradation value: New tools for interpreting multimetric data. *In:* Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making. W.S. Davis and T.P. Simon (ed.). Lewis Publishers. Boca Raton, Florida. (As cited in Karr and Chu 1999).

Appendices

Appendix 1. Recommended Uses of Sediment Quality Guidelines

A1.0 Introduction

Selection of the most appropriate SQGs for specific applications can be a daunting task for sediment assessors. This task is particularly challenging because limited guidance is currently available on the recommended uses of the various SQGs. The following sections provide information on the recommended uses of SQGs in the assessment and management of contaminated sediments. Some of the recommended uses of SQGs at contaminated sites include:

- Designing monitoring programs;
- Interpreting sediment chemistry data;
- Establishing pass/fail levels for dredged material disposal analysis;
- Assessing the risks to biotic receptors associated with contaminated sediments; and,
- Developing site-specific sediment quality remediation objectives.

Each of these uses of SQGs are discussed in the following sections of this appendix.

A1.1 Monitoring Program Design

Monitoring is an integral component of environmental surveillance programs. While such programs may be undertaken for a number of reasons (e.g., trend assessment, impact assessment, compliance), limitations on available resources dictate that they should be conducted in an effective and efficient manner. For this reason, it is important that sediment quality monitoring programs be well focused and provide the type of information that is necessary to manage contaminated sediments.

Sediment quality guidelines contribute to the design of environmental monitoring programs in several ways. First, comparison of existing sediment chemistry data with the SQGs provides a systematic basis for identifying high priority areas for implementing monitoring activities. Second, when used in conjunction with existing sediment chemistry data, the SQGs may be utilized to identify priority contaminants within an area of concern. By considering the potential sources of these contaminants, it may be possible to further identify priority sites for investigation. The SQGs can also assist in monitoring program design by establishing target detection limits for each substance [e.g., threshold effect concentrations (TECs) in MacDonald *et al.* 2000b]. Determination of the detection limits that need to be achieved by analytical laboratories (i.e., to facilitate subsequent interpretation of resultant sediment chemistry data) should help to avoid the difficulties that can result from the use of standard, yet inappropriate, analytical methods (e.g., USEPA contract laboratory procedures; CLP methods).

A1.2 Interpretation of Sediment Chemistry Data

Over the past decade, sediment chemistry data have been collected at a wide range of sites for many purposes. While these data can be used directly to assess the status and trends in environmental quality conditions, they do not, by themselves, provide a basis for determining if the measured concentrations of contaminants represent significant hazards to aquatic organisms. Sediment quality guidelines provide practical assessment tools or "targets" against which the biological significance of sediment chemistry data can be assessed. In this context, SQGs may be used as screening tools to identify areas and COCs (i.e., the substances that are likely to cause or subsequently contribute to adverse biological effects) on site-specific, regional, or national bases.

The numerical SQGs can be used to identify, rank, and prioritize COCs in freshwater, estuarine, and marine sediments. In this application, the concentration of each chemical substance in each sediment sample is compared to the corresponding SQG. Those substances that occur at concentrations below threshold effect-type SQGs (i.e., TECs - MacDonald *et al.* 2000b; TELs - Smith *et al.* 1996; ERLs - USEPA 1996; LELs - Persaud *et al.* 1993; ESGs - USEPA 1997; Appendix 3 of Volume III) should be considered to be of relatively low priority. Those substances that occur at concentrations above the threshold effect-type SQGs but below the probable effect-type SQGs (i.e., PECs - MacDonald *et al.* 2000b; PELs - Smith *et al.* 1996; ERMs - USEPA 1996; SELs - Persaud *et al.* 1993; Appendix 3 of Volume III) should be considered to be of moderate concern, while those that are present at concentrations in excess of the probable effect-type SQGs should be considered to be of relatively high concern. The relative priority that should be assigned to each chemical can be determined by evaluating the magnitude and frequency of exceedance of the SQGs. Chemicals that frequently exceed the probable effect-type SQGs and/or those that exceed the probable effect-type SQGs by large margins should be viewed as the contaminants of greatest concern (Long and MacDonald 1998; MacDonald *et al.* 2000a; 2000b; USEPA 2000b; Ingersoll *et al.* 2001a; 2001b).

In conducting such assessments, it is important to remember that certain chemicals can be present in relatively unavailable forms (such as slag, paint chips, tar). Therefore, there is not a 100% certainty that samples with chemical concentrations in excess of the probable effect-type concentrations will actually be toxic to sediment-dwelling organisms. Therefore, SQGs should be applied with caution in areas with atypical sediment characteristics. Additionally, the reliability of the SQGs should also be considered when conducting evaluations of COCs, with the greatest weight assigned to those SQGs which have been shown to be highly or moderately reliable (USEPA 1996; 2000b; MacDonald *et al.* 2000a; 2000b).

The degree of confidence that can be placed in determinations of COCs can be increased by collecting ancillary sediment quality information. Specifically, data on regional background concentrations of sediment-associated contaminants can be used to identify substances of relatively low concern with respect to anthropogenic activities (i.e., those substances that occur at or below background levels; Appendix 2 of Volume III). Data from toxicity tests can also be used to support the identification of COCs. In particular, matching sediment chemistry and toxicity data provides a basis for evaluating the degree of concordance between the concentrations of specific contaminants and measured adverse effects (USEPA

2000b). The degree of concordance between chemical concentrations and sediment toxicity can be evaluated using correlation analyses and regression plots (Carr *et al.* 1996). Those substances that are present at elevated concentrations (i.e., as indicated by exceedances of the probable effect-type SQGs) in toxic samples should be identified as the contaminants of highest concern (Long and MacDonald 1998; MacDonald *et al.* 2000b). Those chemicals that are not positively correlated with the results of the toxicity tests should be viewed as relatively lower priority.

The numerical SQGs can also be used to identify sites of potential concern with respect to the potential for observing adverse biological effects (Landrum 1995). In this application, the concentrations of sediment-associated contaminants should be compared to the corresponding SQGs. Sediments in which none of the measured chemical concentrations exceed the threshold effect-type SQGs should be considered to have the lowest potential for adversely affecting sediment-dwelling organisms and could be considered as reference areas (Long and Wilson 1997). However, the potential for unmeasured contaminants to be present at levels of toxicological concern can not be dismissed without detailed information on land and water uses within the water body and/or the results of toxicity tests. Those sediments which have concentrations of one of more contaminants between the threshold effect-type SQGs and the probable effect-type SQGs should be considered to be of moderate priority, while those with contaminant concentrations in excess of one or more of the probable effect-type SQGs should be considered to be of relatively high concern. Once again, the magnitude and frequency of exceedances of the probable effect-type SQGs provide a basis for assigned relative priority to areas of concern with respect to contaminated sediments.

While previous guidance has cautioned against using the SQGs as stand alone decision tools, the results of recent evaluations of reliability and predictive ability substantially increase the level of confidence that can be placed in the SQGs. For example, a large database of matching sediment chemistry and toxicity data has been compiled to support an evaluation of the predictive ability of the consensus-based SQGs (USEPA 2000b). The results of this evaluation demonstrated that these consensus-based SQGs provide an accurate basis for classifying sediment samples as toxic or non-toxic, based on bulk sediment chemistry data alone. In this evaluation, mean PEC quotients (PEC-Qs; which provides a measure of overall sediment chemistry relative to the PECs; (USEPA 2000b; Ingersoll *et al.* 2001a) were calculated and used as the primary measure of sediment chemistry. The results of this assessment demonstrated that the incidence of toxicity increased consistently and markedly with increasing mean PEC-Qs (Table A1.1).

Importantly, analysis of the underlying data supported the determination of relationships between mean PEC-Qs and the incidence of toxicity, such that the probability of observing toxicity in any sediment sample can be predicted based on the measured concentrations of trace metals, PAHs, and PCBs. Using these relationships, it was determined that a 50% probability of observing acute and chronic toxicity to the amphipods, *Hyalella azteca*, occurred at mean PEC-Qs of 3.4 and 0.63, respectively (Figure 3.1). Therefore, the probable effect-type SQGs can also be used directly to support certain sediment management decisions, at relatively small sites, where the costs of further investigations could approach the costs of implementing the remedial measures. More costly decisions should be made using multiple lines of evidence to assess sediment quality conditions, however.

Importantly, numerical SQGs provide consistent tools for evaluating spatial patterns in chemical contamination. More specifically, the SQGs can be used to compare and rank sediment quality conditions among basins, waterways, or regions (Long and MacDonald 1998). If a stratified random sampling design is used in the monitoring program, then the SQGs provide a basis for calculating the spatial extent of potentially toxic sediments. In the areas of concern, further investigations would typically be implemented to identify contaminant sources, assess the areal extent and severity of sediment toxicity, evaluate the potential for bioaccumulation, and/or determine the need for source control measures or other remedial measures. The SQGs in combination with sediment chemistry data (Chapter 2 of Volume III), sediment toxicity tests (Chapter 3 of Volume III), benthic invertebrate surveys (Chapter 4 of Volume III), bioaccumulation assessments (Chapter 5 of Volume III), and fish health and fish community assessments (Chapter 6 of Volume III) can also be used to evaluate the success of regulatory actions that are implemented at the site.

A1.3 Analysis of Dredged Materials for Open-Water Disposal

In many waterways, navigational dredging is required to maintain and enhance deep-water harbors and shipping channels. However, questions about the most appropriate means of disposing such dredged materials invariably arise during the planning and implementation of such dredging programs. In the United States, decisions regarding the disposal of dredged materials in freshwater ecosystems are guided by the tiered evaluation process described in the Inland Testing Manual (USEPA and USACE 1998b). Similar guidance has been developed in Canada to assist those involved in navigational dredging and other dredging programs (Porebski 1999). As the Canadian system relies, to a large extent, on SQGs, it provides useful information on the potential applications of numerical SQGs in dredged material assessments.

In Canada, a tiered testing approach has been established to inform decisions regarding the disposal of dredged materials. Using this approach, sediments are considered to be acceptable for open water disposal (for suitable materials in compliance with permit conditions) or beneficial use (e.g., fill, beach nourishment) if the concentrations of all measured contaminants are below screening levels (i.e., threshold effect-type SQGs; TELs). In contrast, sediments are considered to have a high potential for adverse biological effects when the concentrations of one or more contaminants exceed rejection levels (i.e., probable effect-type SQGs; PELs). Such sediments are considered to be unsuitable for open water disposal or for beneficial use.

This tiered approach recognizes that there is a higher level of uncertainty when contaminant concentrations fall between the two guideline levels (i.e., screening and rejection levels). For this reason, sediments with intermediate concentrations of contaminants must undergo biological testing to evaluate their suitability for open water disposal. The biological testing includes a suite of toxicity tests. The applicability of this type of tiered approach is supported by the results of several studies which show that there is a high probability of correctly classifying sediment samples as toxic and not toxic using the SQGs (MacDonald *et al.* 1996; Long *et al.* 1998a; 1998b; MacDonald *et al.* 2000a; 2000b; USEPA 2000b; Ingersoll *et al.* 2001a).

A1.4 Ecological Risk Assessment

Risk assessment is the process of determining the likelihood that adverse effects will occur to ecological receptors in association with exposure to environmental contamination or other hazards. Ecological risk assessment is an evolving process that is designed to provide science-based guidance for managing environmental quality, particularly at contaminated sites. Until recently, appropriate scientific information was not available for assessing the ecological risks that were associated with contaminated sediments. However, a panel of environmental chemists and toxicologists recently concluded that there is sufficient certainty associated with SQGs to recommend their use in ecological risk assessments (Ingersoll *et al.* 1997).

The SQGs contribute directly to several stages of the ecological risk assessment process, including problem formulation, effects assessment, and risk characterization. During problem formulation, background information and Phase I sampling data are used to identify the problem and define the issues that need to be addressed at sites with contaminated sediments (Chapman *et al.* 1997). At the problem formulation stage, SQGs can be used in conjunction with existing sediment chemistry data to identify the chemicals and areas of potential concern with respect to sediment contamination (Long and MacDonald 1998). In turn, this information can be used to scope out the nature and extent of the problem and to identify probable sources of sediment contamination at the site. In addition, the SQGs provide a consistent basis for identifying appropriate reference areas that can be used in subsequent assessments of the site with contaminated sediments (Menzie 1997). Furthermore, the data underlying the SQGs provide a scientific basis for identifying appropriate assessment endpoints (i.e., receptors and function to be protected) and measurement endpoints (i.e., metrics for the assessment endpoints) that can be used at subsequent stages of the assessment.

Numerical SQGs also represent effective tools that can be used to assess the effects of sediment-associated contaminants (i.e., during the effects assessment of an ecological risk assessment). The goal of the effects assessment is to provide information on the toxicity or other effects that are likely to occur in response to exposure to contaminated sediments. In this application, the SQGs provide an effective basis for classifying sediments as toxic or not toxic when used in conjunction with sediment chemistry data (MacDonald *et al.* 1996; USEPA 1996; MacDonald *et al.* 2000b; USEPA 2000b; Ingersoll *et al.* 2001a; 2001b). The applicability of the SQGs in effects assessments is increased when used in conjunction with other tools that facilitate determinations of background concentrations of contaminants, sediment toxicity, bioaccumulation, and effects on *in situ* benthic macroinvertebrates (Chapman *et al.* 1997; Chapter 7 of Volume III).

The primary purpose of the risk characterization stage of an ecological risk assessment is to estimate the nature and extent of the ecological risks at a site with contaminated sediments and to evaluate the level of uncertainty associated with that estimate (Chapman *et al.* 1997). The SQGs are particularly useful at this stage of the process because they provide a quantitative basis for evaluating the potential for observing adverse effects in contaminated sediments, for determining the spatial extent of unacceptable levels of sediment contamination (i.e., sediments that exceed prescribed limits of risk to sediment-dwelling organisms), and for estimating the uncertainty in the risk determinations (i.e., the potential for Type I and Type II errors). Importantly, calculation of the frequency of exceedance of the probable effect-type SQGs and mean SQG quotients for individual sediment samples enables risk assessors to estimate the probability that contaminated sediments will be toxic to sediment-dwelling organisms (Long *et al.* 1998a; 1998b; Field *et al.* 1999; USEPA 2000b). These procedures facilitate determination of the cumulative effects of contaminants arising from multiple sources (i.e., in addition to the contaminated site) and evaluation of the potential for off-site impacts when appropriate sediment chemistry data are available. The uncertainty associated with the application of the guidelines at this stage of the ecological risk assessment can be effectively reduced by using the sediment chemistry data and SQGs in conjunction with other measurement endpoints, such as results of toxicity tests and benthic invertebrate community assessments. Uncertainty associated with establishing cause and effect relationships between SQGs and observed toxicity can be reduced by conducting spiked-sediment exposures and TIE procedures on sediment samples (Ingersoll *et al.* 1997).

A1.5 Development of Sediment Quality Remediation Objectives

Sediment quality remediation objectives (SQROs) are an essential component of the contaminated sediment remediation process because they establish the target clean-up levels for a site. Sediment quality issues are rarely entirely the responsibility of one agency or one level of government. For this reason, it may be necessary to establish agreements between various levels of government to define their respective responsibilities with respect to the prevention, assessment, and remediation of sediment contamination. Multi-jurisdictional agreements may include accords on a number of issues; however, establishment of site-specific SQROs is particularly important because they provide a common yardstick against which the success of a range of sediment management initiatives can be measured (MacDonald and Macfarlane 1999; Ingersoll and MacDonald 1999; MacDonald and Ingersoll 2000).

Numerical SQGs can be used in several ways to support the derivation of SQROs (i.e., clean-up targets). Specifically, SQGs are useful because they provide a means of establishing SQROs that fulfill the narrative use protection objectives for the site (i.e., sediment management objectives). For example, SQROs could be set at chronic effects thresholds if the site management goal is to provide a high level of protection for sediment-dwelling organisms (i.e., PEC-Q of 0.63; USEPA 2000b). Alternatively, the SQROs could be set at acute effects thresholds if the immediate goal for the site is to reduce the potential for acute toxicity and permit natural recovery processes to further reduce risks to sediment-dwelling organisms (i.e., PEC-Q of 3.4; USEPA 2000b). In addition, the SQGs and associated evaluations of predictive ability provide information that may be used to evaluate the costs and benefits associated with various remediation options. Costs-benefit analyses can be further supported by the results of predictive ability analyses, which provide a means of determining the probability of observing adverse effects at various concentrations of sediment-associated contaminants (Field *et al.* 1999; MacDonald *et al.* 2000b; USEPA 2000b).

It is important to note that numerical SQGs should not be regarded as blanket values for regional sediment quality. Variations in environmental conditions among sites could affect sediment quality in different ways and, hence, necessitate the modification of the guidelines to reflect local conditions. MacDonald and Sobolewski (1993) provided interim guidance on the development of site-specific SQROs. In addition, the results of sediment quality triad investigations at the site under investigation can be used to evaluate the applicability of numerical SQGs and to refine these SQGs to make them more directly applicable to the site, if necessary. MacDonald and Ingersoll (2000) provided detailed information on the design and implementation of triad investigations for assessing the predictive ability of SQGs (see also Chapter 7 of Volume III).

Importantly, the weight-of-evidence generated should be proportional to the weight of the decision in the management of contaminated sediments. At small and uncomplicated sites, the costs associated with detailed site investigations are likely to exceed the costs associated with the removal and disposal of contaminated sediments. In these cases, SQGs represent cost-effective tools for establishing clean-up targets and developing remedial action plans. At larger, more complicated sites, it is prudent to conduct further investigations when preliminary screening indicate that contaminated sediments are present. In such cases, the application of toxicity testing, benthic macroinvertebrate community assessments, and other tools provide a means of confirming the severity and extent of degraded sediment quality conditions. Application of TIE procedures and/or sediment spiking studies provides a basis of confirming the identity of the substances that are causing or substantially-contributing to sediment toxicity. In this way, it is possible to design remediation action plans (RAPs) that are most likely to achieve the desired outcomes at the site (i.e., restoration of beneficial uses).

Appendix 2. Methods for Determining Background Levels of Sediment-Associated Contaminants

A2.0 Introduction

Sediment chemistry data is essential for evaluating sediment quality conditions. However, interpretation of environmental data is made difficult by the fact that the measured concentrations of sediment-associated contaminants can be elevated, even in the absence of point source contaminant releases. In some cases, for example, the combination of ambient sediment mineralogy and grain size can result in elevated concentrations of certain metals (Schropp *et al.* 1990; Loring 1991). In addition, the levels of PAHs and other petroleum hydrocarbons can be elevated in the vicinity of naturally-occurring of oil seeps (MacDonald 1994c). Likewise, natural phenomena such as volcanoes and forest fires can release PCDDs and PCDFs into the atmosphere and, ultimately, result in the contamination of sediments (MacDonald 1993). Finally, anthropogenic activities (such as pesticide application or disposal of persistent organic substances) conducted in areas far-removed from the site under consideration can result in elevated levels of PCBs, organochlorine pesticides, and other substances in sediments (i.e., through long-range atmospheric transport and subsequent deposition in aquatic ecosystems (MacDonald 1995). As such, information on contemporary background levels of contaminants in an area is relevant for assessing sediment quality conditions and assessing and remedial options that may be proposed for a site.

The concentrations of trace metals in sediments are influenced by a variety of factors, including sediment mineralogy, grain size, organic content, and anthropogenic enrichment (Schropp and Windom 1988). This combination of factors results in metals levels that can vary over several orders of magnitude at uncontaminated sites (Schropp *et al.* 1990). Therefore, it is important to consider the natural background levels of sediment-associated metals when conducting sediment quality assessments, particularly in regions that have rivers draining metal-rich geologic formations.

There are several procedures available for determining contemporary background levels of contaminants in sediments. In general, these procedures can be grouped into two main categories, including:

- Reference sediment approach; and,
- Reference element approach.

Overviews of these methods for determining contemporary background levels of sediment-associated contaminants are provided in the following sections of this appendix.

A2.1 Reference Sediment Approach

The reference sediment approach involves the determination of regional background levels of metals and/or organic contaminants in the area or region under consideration. Data on regional background levels is important because it provides the information needed to establish contemporary levels of sediment-associated contaminants (i.e., which includes the contribution of chemicals that are associated with human activities, both regionally and at larger geographic scales). One such procedure involves the collection and analysis of surficial sediments from a number of uncontaminated reference sites (i.e., locations that are not affected by known localized contaminant sources) to establish contemporary background concentrations of trace metals or other substances on a regional basis (Persaud *et al.* 1989). In this case, the 95% confidence interval may be used to define the normal range of contaminant concentrations for the region (Reynoldson *et al.* 1995). The upper limit of normal levels can be determined directly from this distribution (i.e., the upper 95% confidence limit; Dunn 1989). Alternatively, the mean plus four standard deviations (i.e., the upper 99% confidence limit) can be used to estimate the upper limit of contemporary background concentrations for the region (IDEM 1992; Adams 1995).

The reference sediment approach can also be used to estimate historic concentrations of trace metals or organic contaminants on a site-specific basis. In this case, sediment coring procedures are used to obtain samples of site sediment from various depths. It is important to collect these cores from fine-grain sediments that have not been disturbed by physical mixing or bioturbation. Chemical analysis of the sub-sections, in conjunction with radiometric dating methods (i.e., ^{137}Cs, ^{210}Pb, or ^{228}Th dating; Valette-Silver 1993; Mudroch and Azcue 1995), provides information for determining how the concentrations of each substance have varied over time. In this way, it is possible to establish the levels of trace metals that correspond to relevant dates in the development of the watershed (i.e., back to the early 1800s). It may be difficult to determine pre-industrial levels of metals if sedimentation rates are high, however (Alexander 1993). Therefore, use of a large-scale regional data base may help provide metal concentrations as a background reference. Statistical methods can be applied to the data that are generated from multiple cores to establish the normal range of background levels for the site under investigation (Reynoldson *et al.* 1995). The upper limit of background can then be established directly from these summary statistics.

A2.2 Reference Element Approach

The reference element approach was developed to provide a basis for assessing metal contamination in sediments (Loring 1991; Schropp and Windom 1988; Schropp *et al.* 1990; Schiff and Weisberg 1996). This procedure relies on normalization of metal concentrations to a reference element. Normalization of metal concentrations to concentrations of aluminum in estuarine sediments provided the most useful method of comparing metal levels on a regional basis in Florida estuaries. However, normalization using lithium, iron, or other reference elements has been used in other estuarine regions (Loring 1991; Schiff and Weisberg 1996). Recently, Carvalho and Schropp (2001) demonstrated that normalization of metal concentrations to the concentrations of aluminum also provides an effective basis for evaluating metal enrichment in freshwater sediments.

Development of the metals interpretive tool is a relatively straight forward process. Briefly, data on sediment metal concentrations are collected from roughly 100 sites chosen for being remote from known or potential sources of metals contamination. Total metal concentrations are determined in each of these samples. Simple linear regressions of the concentrations of each of seven metals to aluminum concentrations are performed on log-transformed data and 95% prediction limits are calculated. The regression lines and prediction limits are then plotted. These plots then form the basis for interpreting data on the concentrations of metals in sediments, such that anthropogenic enrichment of metal levels would be suspected at sites with metals concentrations exceeding the upper 95% prediction limit (for one or more substances). The application of this procedure using data from various estuarine areas (e.g., Tampa Bay, Schropp *et al.* 1989; Louisiana, Pardue *et al.* 1992) has supported the effectiveness and utility of this interpretive tool. A comparable tool for assessing metal enrichment in freshwater sediments has been developed for the State of Florida (Figure A2-1; Carvalho and Schropp 2001).

Appendix 3. Approaches to the Development of Numerical Sediment Quality Guidelines

A3.0 Introduction

Numerical SQGs (including ESGs, sediment quality objectives, and sediment quality standards) have been developed by various jurisdictions in North America for both freshwater and marine ecosystems. The SQGs that are currently being used in North America have been developed using a variety of approaches, including both empirical and theoretical approaches. Both empirical and theoretical approaches were considered to support the derivation numerical SQGs for the protection of sediment-dwelling organisms, including:

- Screening Level Concentration Approach (SLCA);
- Effects Range Approach (ERA);
- Effects Level Approach (ELA);
- Apparent Effects Threshold Approach (AETA);
- Equilibrium Partitioning Approach (EqPA).
- Logistic Regression Modeling Approach (LRMA); and,
- Consensus Approach (CA).

The tissue residue approach represents the primary method for deriving numerical SQGs for the protection of wildlife and human health (i.e., for substances that bioaccumulate in the food web). The following sections of this report provide brief descriptions of each of these approaches.

A3.1 Screening Level Concentration Approach

The screening level concentration approach (SLCA) is a biological effects-based approach that is applicable to the development of SQGs for the protection of benthic organisms. This approach utilizes matching biological and chemistry data collected in field surveys to calculate a screening level concentration (SLC; Neff *et al.* 1986). The SLC is an estimate of the highest concentration of a contaminant that can be tolerated by a pre-defined proportion of benthic infaunal species.

The SLC is determined through the use of a database that contains information on the concentrations of specific contaminants in sediments and on the co-occurrence of benthic organisms in the same sediments. For each benthic organism for which adequate data are available, a species screening level concentration (SSLC) is calculated. The SSLC is determined by plotting the frequency distribution of the contaminant concentrations over all of the sites at which the species occurs (information from at least ten sites is required to calculate a SSLC). The 90th percentile of this distribution is taken as the SSLC for the species being investigated. The SSLCs for all of the species for which adequate data are available are then compiled as a frequency distribution to determine the concentration that can be tolerated by a specific proportion of the species (i.e., the 5th percentile of the distribution would provide an SLC that should be tolerated by 95% of the species). This concentration is termed the screening level concentration of the contaminant.

A number of jurisdictions have used the SLCA to derive numerical SQGs. In the St. Lawrence River, two SQGs were developed for five groups of PCBs using the SLCA, including a minimal effect threshold (MET) and a toxic effect threshold (TET; EC and MENVIQ 1992). The MET was calculated as the 15th percentile of the SSLCs, while the TET was calculated as the 90th percentile of the SSLC distribution for each substance. Therefore, the MET and TET are considered to provide protection for 85% and 10% of the species represented in the database, respectively. Similarly, Environment Ontario developed a lowest effect level (LEL) and severe effect level (SEL) using this approach (Persaud *et al.* 1993). Neff *et al.* (1986) also developed a screening level concentration (SLC) for tPCBs primarily using data from the Great Lakes.

A3.2 Effects Range Approach

The effects range approach (ERA) to the derivation of SQGs was developed to provide informal tools for assessing the potential for various contaminants tested in the National Status and Trends Program (NSTP) to be associated with adverse effects on sediment-dwelling organisms (Long and Morgan 1991). As a first step, a database was compiled which contained information on the effects of sediment-associated contaminants, including data from spiked-sediment toxicity tests, matching sediment chemistry and biological effects data from field studies in the United States, and SQGs that were derived using various approaches. All of the information in the database was weighted equally, regardless of the method that was used to develop it. The objective of this initiative was to identify informal guidelines which could be used to evaluate sediment chemistry data collected nationwide under the NSTP.

Candidate data sets from field studies were evaluated to determine their applicability for incorporation into the database. This evaluation was designed to determine the overall applicability of the data set, the methods that were used, the end-points that were measured, and the degree of concordance between the chemical and biological data. The data which met the evaluation criteria were incorporated into the database (Long and Morgan 1991; Long *et al.* 1995).

The database that was compiled included several types of information from each study. Individual entries consisted of the concentration of the contaminant, the location of the study, the species tested and endpoint measured, and an indication of whether or not there was concordance between the observed effect and the concentrations of a specific chemical (i.e., no effect, no or small gradient, no concordance, or a "hit", which indicated that an effect was measured in association with elevated sediment chemistry). Data from non-toxic or unaffected samples were assumed to represent background conditions. Data which showed no concordance between chemical and biological variables were included in the database, but were not used to calculate the SQGs. The data for which a biological effect was observed in association with elevated chemical concentrations (i.e., hits) were sorted in ascending order of concentration and the 10th and 50th percentile concentrations for each compound were determined. The effects range-low (ERL; 10th percentile value) was considered to represent a lower threshold value, below which adverse effects on sensitive life stages and/or species occurred infrequently. The effects range-median (ERM; 50th percentile value) was considered to represent a second threshold value, above which adverse effects were frequently observed. These two parameters, ERL and ERM, were then used as informal SQGs (Long and Morgan 1991; Long *et al.* 1995). USEPA (1996) used a similar approach to derive ERLs (15th percentile of the effects data set) and ERMs (50th percentile of the effects data set) for assessing sediments from various freshwater locations. Similarly, MacDonald (1997) applied the effects range approach to regionally-collected field data to derive site-specific sediment effect concentrations for PCBs and DDTs in the Southern California Bight.

A3.3 Effects Level Approach

The effects level approach (ELA) is closely related to the effects range approach described above. However, the ELA is supported by an expanded version of the database that was used to derive the effects levels (Long and Morgan 1991). The expanded database contains matching sediment chemistry and biological effects data from spiked-sediment toxicity tests and from field studies conducted throughout North America (including both effects and no effects data). The expanded database also contains SQGs derived using various approaches. The information contained in the expanded database was evaluated and classified in the same manner that was used to compile the original NSTP database.

In the ELA, the underlying information in the database was used to derive two types of SQGs, including threshold effect levels (TELs) and probable effect levels (PELs). The TEL, which is calculated as the geometric mean of the 15th percentile of the effects data set and the 50th percentile of the no effects data set, represents the chemical concentration below which adverse effects occurred only infrequently. The PEL represents a second threshold value, above which adverse effects were frequently observed. The PEL is calculated as the geometric mean of the 50th percentile of the effects data set and the 85th percentile of the no effects data set. These arithmetic procedures have been applied to the expanded database to derive numerical SQGs (i.e., TELs and PELs) for Florida coastal waters (MacDonald *et* *al.* 1996), United States freshwater systems (USEPA 1996), and Canadian freshwater and marine systems (Smith *et al.* 1996; CCME 1999).

A3.4 Apparent Effects Threshold Approach

The apparent effects threshold approach (AETA) to the development of SQGs was developed for use in the Puget Sound area of Washington State (Tetra Tech Inc. 1986). The AETA is based on empirically-defined relationships between measured concentrations of a contaminant in sediments and observed biological effects. This approach is intended to define the concentration of a contaminant in sediment above which significant ($p \leq 0.05$) biological effects are observed. These biological effects include, but are not limited to, toxicity to benthic and/or water column species (as measured using sediment toxicity tests), changes in the abundance of various benthic species, and changes in benthic community structure. In Puget Sound, for example, four AET values have been generated, including AETs for Microtox, oyster larvae, benthic community, and amphipods. The AET values are based on dry weight-normalized contaminant concentrations for metals and either dry weight- or TOC-normalized concentrations for organic substances (Barrick *et al.* 1988; Washington Department of Ecology 1990). The state of Washington has used the various AET values to establish sediment quality standards and minimum clean-up levels for COCs in the state.

Cubbage *et al.* (1997) refined this approach to support the development of probable AETs (PAETs) using matching sediment chemistry and toxicity data for freshwater sediments from the state of Washington. USEPA (1996) utilized a similar approach to develop freshwater AETs (termed no effect concentrations or NECs in that study) using data from various freshwater locations.

A3.5 Equilibrium Partitioning Approach

The water-sediment equilibrium partitioning approach (EqPA) has been one of the most studied and evaluated approaches for developing SQGs (sometimes termed ESGs) for non-ionic organic chemicals and metals (Pavlou and Weston 1983; Bolton *et al.* 1985; Kadeg *et al.* 1986; Pavlou 1987; Di Toro *et al.* 1991; Ankley *et al.* 1996; Hansen *et al.* 1996). This approach is based on the premise that the distribution of contaminants among different compartments in the sediment matrix (i.e., sediment solids and pore water) is predictable based on their physical and chemical properties, assuming that continuous equilibrium exchange between sediment and pore water occurs. This approach has been supported by the results of spiked-sediment toxicity tests, which indicate that positive correlations exist between the biological effects observed and the concentrations of contaminants measured in the pore water (Di Toro *et al.* 1991; Ankley *et al.* 1996; Berry *et al.* 1996; Hansen *et al.* 1996). A primary strength of the EqPA approach is that the bioavailability of individual classes of compounds (i.e., metals or non-ionic organic compounds) can be addressed.

In the EqPA, water quality criteria developed for the protection of freshwater or marine organisms are used to support the SQGs derivation process. As such, the water quality criteria formulated for the protection of water column species are assumed to be applicable to benthic organisms (Di Toro *et al.* 1991). The ESGs are calculated using the appropriate water quality criteria (usually the final chronic values, FCVs, or equivalent values; USEPA 1997) in conjunction with the sediment/water partition coefficients (Kp) for the specific contaminants (note that other effect concentrations (e.g., an LC_{50} for a particular species of concern) can also be used in the calculation of ESGs). The final chronic value is derived from the species mean chronic values that have been calculated from published toxicity data and is intended to protect 95% of aquatic species. The calculation procedure for non-ionic organic contaminants is as follows:

$$SQG = Kp \bullet FCV$$

where:

SQG = Sediment quality guideline (in µg/kg);
Kp = Partition coefficient for the chemical (in L/kg); and,
FCV = Final Chronic Value (in µg/L).

The Kp is a function of the partition coefficient for sediment organic carbon (K_{oc}) of the substance under consideration and the amount of organic carbon in the sediment under investigation (f_{oc}; where $Kp = K_{oc} \cdot f_{oc}$; Di Toro *et al.* 1991). The K_{oc} for non-ionic substances can be calculated from its K_{ow} (Di Toro *et al.* 1991). Procedures for evaluating the potential for sediment toxicity due to the presence of metals have also been developed (Ankley *et al.* 1996). These procedures rely on the determination of AVS and SEM concentrations. Samples in which the molar concentrations of AVS equal or exceed the molar concentrations of five divalent metals (Cd, Cu, Pb, Ni, Zn) are unlikely to be toxic due to metals. In contrast, samples with SEM-AVS >1 could be toxic due to metals (Ankley *et al.* 1996). Based on the results of more recent analyses, SEM-AVS >5 may be a better predictor of toxicity due to the presence of divalent cationic metals.

A3.6 Logistic Regression Modeling Approach

In the logistic regression modeling approach (LRMA), numerical SQGs are derived from the results of field studies of sediment quality conditions in marine and estuarine habitats. The first step in this process involves the collection, evaluation, and compilations of matching sediment chemistry and toxicity data from a wide variety of sites in North America (Field *et al.* 1999; 2001). Next, the information that were compiled in the database were retrieved on a substance-by-substance basis, with the data from individual sediment samples sorted in order of ascending concentration. For each sediment sample, the ascending data table was used to provide information on the concentration of contaminant under consideration (on either a dry weight- or organic carbon-normalized basis) and the toxicity test results (i.e., toxic or not toxic) for each toxicity test endpoint (e.g., 10-day survival of marine amphipods).

In the next step of the process, the data contained in the ascending data tables were screened to minimize the potential for including samples in which the selected contaminant did not contribute substantially to the observed toxicity. In this analysis, the chemical concentration in each toxic sample was compared to the mean concentration in the non-toxic samples from the same study and geographic area. The toxic samples with concentrations of the selected non-toxic samples were not used to develop the models for each COPC (i.e., it was highly unlikely that the contaminant substantially contributed to sediment toxicity in such samples).

In the final step of the analysis, the screened data were used to develop logistic regression models, which express the relationship between the concentration of the selected contaminant and the probability of observing toxicity. In its simplest form, the logistic model can be described using the following equation:

$$p = e^{B0+B1(x)} \div (1 + e^{B0+B1(x)})$$

where:

p = probability of observing a toxic effect;
B0 = intercept parameter;
B1 = slope parameter; and,
x = concentration or log concentration of the chemical.

Using databases consisting of the results of 10-day amphipod toxicity tests, Field *et al.* (1999; 2002) derived logistic regression models for several chemical substances to illustrate the methodology. More specifically, these studies calculated T10, T20, T50, T80, or T90 values for several metals, PAHs, and total PCBs. These values represent the chemical concentrations that correspond to a 10%, 20%, 50%, 80%, or 90% probability of observing sediment toxicity. In addition to supporting the derivation of specific T-values, this method can be used to determine the concentration of a contaminant that corresponds to any probability of observing toxicity. Therefore, a sediment manager can identify an acceptable probability of observing sediment toxicity at a site (e.g., 25%) and determine the corresponding chemical concentrations (e.g., T25 value). The calculated value can then be used as the SQG for the site. This procedure is currently being used to evaluate data as part of a second report to Congress on sediment quality (an update to USEPA 1997).

A3.7 Consensus Approach

In the consensus approach (CA), consensus-based SQGs were derived from the existing SQGs that have been published for the protection of sediment-dwelling organisms (MacDonald *et al.* 2000a; 2000b). Derivation of numerical SQGs using the CA involved as four stepped process. In a first step, the SQGs that have been derived by various investigators for assessing the quality of freshwater sediments were collected and collated. Next, the SQGs obtained from all sources were evaluated to determine their applicability to the derivation of consensus-based SQGs. The selection criteria that were applied are intended to evaluate the transparency of the derivation methods, the degree to which the SQGs are effects-based, and the uniqueness of the SQGs.

The effects-based SQGs that meet these selection criteria were then grouped in MacDonald *et al.* (2000a; 2000b) to facilitate the derivation of consensus-based SQGs (Swartz 1999). Specifically, the SQGs for the protection of sediment-dwelling organisms were grouped into two categories according to their original narrative intent, including TECs and probable effect concentrations (PECs). The TECs were intended to identify contaminant concentrations below which harmful effects on sediment-dwelling organisms were unlikely to be observed. Examples of TEC-type SQGs include threshold effect levels (TELs; Smith *et al.* 1996; USEPA 1996), effect range low values (ERLs; Long and Morgan 1991; USEPA 1996), lowest effect levels (LELs; Persaud *et al.* 1993), and chronic equilibrium partitioning thresholds (USEPA 1997). The PECs were intended to identify contaminant concentrations above which harmful effects on sediment-dwelling organisms were likely to be frequently or always observed (MacDonald *et al.* 1996; Swartz 1999). Examples of PEC-type SQGs include probable effect levels (PELs; Smith *et al.* 1996; USEPA 1996a), effect range median values (ERMs; Long and Morgan 1991; USEPA 1996a); and severe effect levels (Persaud *et al.* 1993).

Following classification of the published SQGs, consensus-based TECs were calculated by determining the geometric mean of the SQGs that were included in this category. Likewise, consensus-based PECs were calculated by determining the geometric mean of the PEC-type values. The geometric mean, rather than the arithmetic mean, was calculated because it provided an estimate of central tendency that was not unduly affected by outliers and because the distributions of the SQGs were not known. Consensus-based TECs or PECs were calculated only if three of more published SQGs are available for a chemical substance or group of substances (MacDonald *et al.* 2000a; 2000b).

The consensus approach has been used to derive numerical SQGs for a variety of chemical substances and media types. For example, Swartz (1999) derived consensus-based SQGs for PAHs in marine ecosystems. More recently, MacDonald *et al.* (2000a) derived SQGs for total PCBs in freshwater and marine sediments. Ingersoll and MacDonald (1999) and MacDonald *et al.* (2000a; 2000b) have also developed consensus-based SQGs for metals, PAHs, PCBs, and several pesticides in freshwater sediments. USEPA (2000b) and Ingersoll *et al.* (2001a) used consensus-based SQGs to evaluate the incidence of toxicity in a national freshwater database. As the term implies, consensus-based SQGs are intended to reflect the agreement among the various SQGs by providing an estimate of their central tendency. Consensus-based SQGs are, therefore, considered to provided a unifying synthesis of the existing SQGs, reflect causal rather than correlative effects, and account for the effects of contaminant mixtures in sediment (Swartz 1999; Di Toro and McGrath 2000; MacDonald *et al.* 2000a; 2000b).

A3.8 Tissue Residue Approach

The tissue residue approach (TRA; which is also known as the biota-water-sediment EqPA) is based on the fact that sediments represent important sources of contaminants that bioaccumulate in the tissues of aquatic organisms and are transferred into aquatic food webs. For this reason, it is necessary to assure that the concentrations of sediment-associated contaminants remain below the levels that are associated with the accumulation of such contaminants to harmful levels in sediment-dwelling organisms and other elements of the food web. Therefore, application of the TRA involves the establishment of safe sediment concentrations for individual chemicals or classes of chemicals by determining the chemical concentrations in sediments that are predicted to result in acceptable tissue residues (i.e., in fish and shellfish tissues that are consumed by piscivorus wildlife).

Derivation of numerical SQGs using the TRA involves several steps. As a first step, the contaminants for which SQGs are to be derived are selected based on their potential to accumulate in aquatic food webs. Next, numerical TRGs are identified for these contaminants. Three types of TRGs may be used to derive the SQGs, including:

- Critical body burdens in sediment-dwelling organisms, which define the threshold levels of tissue-associated contaminants relative to adverse effects on benthic species (e.g., Jarvinen and Ankley 1999);
- Tissue residue guidelines for the protection of aquatic-dependent wildlife, which define tolerable levels of contaminants in fish and aquatic invertebrates that are consumed by avian and mammalian receptors (e.g., Newell *et al.* 1987); and
- Tissue residue guidelines for the protection of human health, which define tolerable levels of contaminants in fish and shellfish that are consumed by humans (e.g., Federal Drug Administration Action Levels).

Following the selection of TRGs, BSAFs are determined each of the substances of concern. Such BSAFs can be determined from the results of bioaccumulation assessments, from matching sediment chemistry and tissue residue data collected in the field, and/or from the results of bioaccumulation models. Such BSAFs must be relevant to the species under consideration (i.e., laboratory-derived BSAFs for polychaetes should not be used directly to estimate BSAFs in fish). Numerical SQGs are subsequently derived using the equation:

$$SQG = TRG \div BSAF$$

This approach has been used on several occasions to develop SQGs for the protection of human health, most frequently for DDTs, mercury, and PCBs. In addition, SQGs for 2,3,7,8 tetrachlorodibenzo-*p*-dioxin (T_4CDD) have been established for Lake Ontario on the basis of fish tissue residues (Endicott *et al.* 1989; Cook *et al.* 1989). The applicability of this approach to the derivation of SQGs is supported by data which demonstrate that declines in DDT residues in fish and birds (since its use was banned) are strongly correlated with declining concentrations of this substance in surficial sediments in the Great Lakes and Southern California Bight. As such, this approach is a logical companion for the EqPA and the other approaches that were described previously. However, uncertainty in the selection of critical body burdens in sediment-dwelling organisms limits the applicability of this approach for deriving SQGs for the protection of benthic invertebrate species.

Appendix 4. Criteria for Evaluating Candidate Data Sets

A4.0 Introduction

In recent years, the Great Lakes National Program Office (USEPA), United States Geological Survey, National Oceanic and Administration, Minnesota Pollution Control Agency, Florida Department of Environmental Protection, British Columbia Ministry of Water, Air and Land Protection, MacDonald Environmental Sciences Ltd., and EVS Consultants have been developing a database of matching sediment chemistry and sediment toxicity data to support evaluations of the predictive ability of numerical SQGs in the Great Lakes Basin and elsewhere in North America (Field *et al.* 1999; USEPA 2000b; Crane *et al.* 2000). In addition, various project-specific databases have been developed to facilitate access to and analysis of data sets to support natural resource damage assessments and ecological risk assessments at sites with contaminated sediments (MacDonald and Ingersoll 2000; Crane *et al.* 2000; MacDonald *et al.* 2001a; 2001b; Ingersoll *et al.* 2001a). The goal of these initiatives was to collect and collate the highest quality data sets for assessing sediment quality conditions at contaminated sites and evaluating numerical SQGs. To assure that the data used in these assessments met the associated DQOs, all of the candidate data sets were critically evaluated before inclusion in the database. However, the screening process was also designed to be flexible to assure that professional judgement could also be used when necessary in the evaluation process. In this way, it was possible to include as many data sets as possible and, subsequently, use them to the extent that the data quality and quantity dictate.

The following criteria for evaluating candidate data sets were established in consultation with an *ad hoc* Science Advisory Group on Sediment Quality Assessment (which is comprised of representatives of federal, provincial, and state government agencies, consulting firms, and non-governmental organizations located throughout North America and elsewhere worldwide). These criteria are reproduced here because they provide useful guidance on the evaluation of data that have been generated to support sediment quality assessments. In addition, these criteria can be used to support the design of sediment sampling and analysis plans, and associated quality assurance project plans (see Volume II).

A4.1 Criteria for Evaluating Whole Sediment, Pore Water, and Tissue Chemistry

Data on the chemical composition of whole sediments, pore water, and biological tissues are of fundamental importance in assessments of sediment quality conditions. For this reason, it is essential to ensure that high quality data are generated and used to support such sediment quality assessments. In this respect, data from individual studies are considered to be acceptable if:

- Samples were collected from any sediment horizon (samples representing surficial sediments are most appropriate for assessing effects on sediment-dwelling organisms and other receptors, while samples of sub-surface sediments are appropriate for assessing potential effects on sediment-dwelling organisms and other receptors, should these sediments become exposed; ASTM 2001a; ASTM 2001e; USEPA 2000a);
- Appropriate procedures were used for collecting, handling, and storing sediments (e.g., ASTM 2001c; 2001d; USEPA 2001) and samples of other media types;
- The concentrations of a variety of COPCs were measured in samples;
- Appropriate analytical methods were used to generate chemistry data. The methods that are considered to be appropriate included USEPA approved methods, other standardized methods (e.g., ASTM methods, SW-846 methods), or methods that have been demonstrated to be equivalent or superior to standard methods; and,
- Data quality objectives were met. The criteria that are used to evaluate data quality included:

  (i) the investigator indicated that DQOs had been met;
  (ii) analytical detection limits were reported and lower than the PECs (however, detection limits < TEC are preferred);
  (iii) accuracy and precision of the chemistry data were reported and within acceptable ranges for the method;

(iv) sample contamination was not noted (i.e., analytes were not detected at unacceptable concentrations in method blanks); and,

(v) the results of a detailed independent review indicated that the data were acceptable and/or professional judgement indicated that the data set was likely to be of sufficient quality to be used in the assessment (i.e., in conjunction with author communications and/or other investigations).

A4.2 Criteria for Evaluating Biological Effects Data

Data on the effects of contaminated sediments on sediment-dwelling organisms and other aquatic species provide important information for evaluating the severity and extent of sediment contamination. Data from individual studies are considered to be acceptable for this purpose if:

- Appropriate procedures were used for collecting, handling, and storing sediments (e.g., ASTM 2001c; USEPA 2000a; 2001); Sediments were not frozen before toxicity tests were initiated (ASTM 2001a; 2001e);

- The responses in the negative control and/or reference groups were within accepted limits (i.e., ASTM 2001a; 2001d; 2001e; 2001f; 2001g; 2001h; USEPA 2000b);

- Adequate environmental conditions were maintained in the test chambers during toxicity testing (i.e., ASTM 2001a; 2001e; USEPA 2000b);

- The endpoint(s) measured were ecologically-relevant (i.e., likely to influence the organism's viability in the field) or indicative of ecologically-relevant endpoints; and,

- Appropriate procedures were used to conduct bioaccumulation tests (ASTM 2001d).

Additional guidance is presented in USEPA (1994) and in Chapter 4 of Volume III for evaluating the quality of benthic community data generated as part of a sediment quality assessment. These criteria include collection of replicate samples, resorting at least 10% of the samples, and independent checks of taxonomic identification of specimens. Guidance is presented in USEPA (2000c) and in Schmidt *et al.* (2000) for evaluating the quality of fish health and fish community data.

Tables

Table 2.1. Sediment quality guidelines that reflect threshold effect concentrations (TECs; i.e., below which harmful effects are unlikely to be observed; from MacDonald *et al.* 2000b).

Substance	*Threshold Effect Concentrations*						
	TEL	**LEL**	**MET**	**ERL**	**TEL-HA28**	**SQAL**	**Consensus-Based TEC**
Metals (in mg/kg DW)							
Arsenic	5.9	6	7	33	11	NG	9.79
Cadmium	0.596	0.6	0.9	5	0.58	NG	0.99
Chromium	37.3	26	55	80	36	NG	43.4
Copper	35.7	16	28	70	28	NG	31.6
Lead	35	31	42	35	37	NG	35.8
Mercury	0.174	0.2	0.2	0.15	NG	NG	0.18
Nickel	18	16	35	30	20	NG	22.7
Zinc	123	120	150	120	98	NG	121
Polycyclic Aromatic Hydrocarbons (PAHs; in µg/kg DW)							
Anthracene	NG	220	NG	85	10	NG	57.2
Fluorene	NG	190	NG	35	10	540	77.4
Naphthalene	NG	NG	400	340	15	470	176
Phenanthrene	41.9	560	400	225	19	1800	204
Benz[a]anthracene	31.7	320	400	230	16	NG	108
Benzo(a)pyrene	31.9	370	500	400	32	NG	150
Chrysene	57.1	340	600	400	27	NG	166
Dibenz[a,h]anthracene	NG	60	NG	60	10	NG	33.0
Fluoranthene	111	750	600	600	31	6200	423
Pyrene	53	490	700	350	44	NG	195
Total PAHs	NG	4000	NG	4000	260	NG	1610

Table 2.1. Sediment quality guidelines that reflect threshold effect concentrations (TECs; i.e., below which harmful effects are unlikely to be observed; from MacDonald *et al.* 2000b).

Substance	*Threshold Effect Concentrations*						
	TEL	**LEL**	**MET**	**ERL**	**TEL-HA28**	**SQAL**	**Consensus-Based TEC**
Polychlorinated Biphenyls (PCBs; in µg/kg DW)							
Total PCBs	34.1	70	200	50	32	NG	59.8
Organochlorine Pesticides (in µg/kg DW)							
Chlordane	4.5	7	7	0.5	NG	NG	3.24
Dieldrin	2.85	2	2	0.02	NG	110	1.90
Sum DDD	3.54	8	10	2	NG	NG	4.88
Sum DDE	1.42	5	7	2	NG	NG	3.16
Sum DDT	NG	8	9	1	NG	NG	4.16
Total DDTs	7	7	NG	3	NG	NG	5.28
Endrin	2.67	3	8	0.02	NG	42	2.22
Heptachlor epoxide	0.6	5	5	NG	NG	NG	2.47
Lindane (gamma-BHC)	0.94	3	3	NG	NG	3.7	2.37

TEC = Threshold effect concentration (from MacDonald *et al.* 2000a).
TEL = Threshold effect level; dry weight (Smith *et al.* 1996).
LEL = Lowest effect level, dry weight (Persaud *et al.* 1993).
MET = Minimal effect threshold; dry weight (EC & MENVIQ 1992).
ERL = Effects range low; dry weight (Long and Morgan 1991).
TEL-HA28 = Threshold effect level for *Hyalella azteca*; 28 day test; dry weight (USEPA 1996).
SQAL = Sediment quality advisory levels; dry weight at 1% OC (USEPA 1997).
NG = No guideline; DW = dry weight.

Table 2.2. Sediment quality guidelines that reflect probable effect concentrations (PECs; i.e., above which harmful effects are likely to be observed; from MacDonald *et al.* 2000b).

	Probable Effect Concentrations					
Substance	**PEL**	**SEL**	**TET**	**ERM**	**PEL-HA28**	**Consensus-Based PEC**
Metals (in mg/kg DW)						
Arsenic	17	33	17	85	48	33.0
Cadmium	3.53	10	3	9	3.2	4.98
Chromium	90	110	100	145	120	111
Copper	197	110	86	390	100	149
Lead	91.3	250	170	110	82	128
Mercury	0.486	2	1	1.3	NG	1.06
Nickel	36	75	61	50	33	48.6
Zinc	315	820	540	270	540	459
Polycyclic Aromatic Hydrocarbons (PAHs; in μg/kg DW)						
Anthracene	NG	3700	NG	960	170	845
Fluorene	NG	1600	NG	640	150	536
Naphthalene	NG	NG	600	2100	140	561
Phenanthrene	515	9500	800	1380	410	1170
Benz[a]anthracene	385	14800	500	1600	280	1050
Benzo(a)pyrene	782	14400	700	2500	320	1450
Chrysene	862	4600	800	2800	410	1290
Fluoranthene	2355	10200	2000	3600	320	2230
Pyrene	875	8500	1000	2200	490	1520
Total PAHs	NG	100000	NG	35000	3400	22800
Polychlorinated Biphenyls (PCBs; in μg/kg DW)						
Total PCBs	277	5300	1000	400	240	676

Table 2.2. Sediment quality guidelines that reflect probable effect concentrations (PECs; i.e., above which harmful effects are likely to be observed; from MacDonald *et al.* 2000b).

Substance	PEL	SEL	TET	ERM	PEL-HA28	Consensus-Based PEC
			Probable Effect Concentrations			
Organochlorine Pesticides (in µg/kg DW)						
Chlordane	8.9	60	30	6	NG	17.6
Dieldrin	6.67	910	300	8	NG	61.8
Sum DDD	8.51	60	60	20	NG	28.0
Sum DDE	6.75	190	50	15	NG	31.3
Sum DDT	NG	710	50	7	NG	62.9
Total DDTs	4450	120	NG	350	NG	572
Endrin	62.4	1300	500	45	NG	207
Heptachlor Epoxide	2.74	50	30	NG	NG	16.0
Lindane (gamma-BHC)	1.38	10	9	NG	NG	4.99

PECs = probable effect concentrations (from MacDonald *et al.* 2000a)
PEL = Probable effect level; dry weight (Smith *et al.* 1996).
SEL = Severe effect level, dry weight (Persaud *et al.* 1993).
TET = Toxic effect threshold; dry weight (EC & MENVIQ 1992).
ERM = Effects range median; dry weight (Long and Morgan 1991).
PEL-HA28 = Probable effect level for *Hyalella azteca* ; 28-day test; dry weight (USEPA 1996a).
NG = No guideline; DW = dry weight.

Table 2.3. Advantages and disadvantages of whole sediment and pore water chemistry (Ingersoll *et al.* 1997).

Advantages	Disadvantages
• Provides direct information for determining the presence/absence of COPCs.	• Can not be used to evaluate effects on ecological receptors directly.
• Standard methods are available for most COPCs.	• Effective interpretation of the data is dependent on selecting the appropriate suite of analytes.
• Procedures are available for evaluating the reliability of the data (i.e., accuracy and precision).	• The use of inappropriate methods (e.g., with high detection limits) can limit the utility of the resultant data.
• Methods for assessing the bioavailability of COPCs are available.	• For pore water, it is challenging to obtain sufficient sample volumes to support the desired chemical analysis.
• Benchmarks (i.e., SQGs) are available for many COPCs for evaluating the potential for biological effects.	• Pore water extraction methods can alter pore water chemistry.

Table 2.4. Uncertainty associated with sediment chemistry measurements (Ingersoll *et al.* 1997).

	Bulk Sediment	Total Organic Carbon Normalization	SEM minus AVS	Metal Speciation (non AVS)	Pore water	Elutriate	Reference Element
Precision	1	1	1	2*	2*	1	1
Ecological relevance	3	2	2	2	2	3	3
Causality: Contaminant	1	1	1	1	2	3	1
Causality: Source	2*	2	2	2	2	3	1
Sensitivity	1*	1*	1*	1*	1*	1*	1*
Interference	2*	2*	2*	2*	2*	2*	2*
Standardization	1*	1*	1*	3*	2*	2*	1*
Discrimination	1	1	1	2*	1	1	1
Bioavailability	2*	1	1*	2*	2*	3	2
Field validation[a]	1	2	2*	2*	3	3	1

Ranking Code: 1 = low uncertainty (good); 3 = high (bad); * = lack of knowledge.

[a] Not related to field sampling.

Table 2.5. Uncertainty associated with sediment quality guidelines (Ingersoll *et al.* 1997).

	ESGs[1]	ERL and ERM	AET	SLC	SEM-AVS	Toxic Unit Models	Residue-Based SQG
Precision	1	2*	3	3	2*	3*	2
Ecological relevance	2*	1	2	3	1	1	2*[b]
Causality	1	3*a	3	3	1	1*	2
Sensitivity	2	1[c]/2	3	1	2*	2	1
Interference[d]	2	2*	2	3	2*	2	1
Standardization	1	1	2	2	1	3	1
Discrimination[e]	1	1	3	3	2	2	1
Bioavailability	1	2*[e]	2*[e]	1	1	2*	1*
Field validation	2*	2*	2*	3*	2*	2	2*

Ranking Code: 1 = low uncertainty (good); 3 = high (bad); * = lack of knowledge.

[a] with TU; [b] few compounds, based on consumption effects; [c] ERL; [d] interferences resulting from community responses and mixture effects; [e] with normalization.

[1] ESG = Equilibrium-Based Sediment Guidelines (formerly known as Sediment Quality Criteria in Ingersoll *et al.* 1997).

Table 2.6. Summary of potential targets for pore water chemistry.

Analyte	WQ Criteria		Reference	LC$_{50}$ for *Hyalella azteca*	Reference	Invertebrates		Reference
	Acute	Chronic				Acute	Chronic	
Metals								
Aluminum								
Arsenic	340 µg/L	150 µg/L	USEPA 1999					
Cadmium	4.3 µg/L	2.2 µg/L	USEPA 1999	2.94 µg/L	USEPA 1994	3.6[6]	0.17[6]	Outridge *et al.* 1994
Chromium						15[2]	2.5[2]	CCREM 1987
Chromium (III)	570 µg/L	74 µg/L	USEPA 1999					
Chromium (VI)	16 µg/L	11 µg/L	USEPA 1999					
Copper	13 µg/L	9 µg/L	USEPA 1999	35 µg/L	USEPA 1994	20[1]	8[1]	Spear and Pierce 1979
Lead	65 µg/L	2.5 µg/L	USEPA 1999	< 16 µg/L	USEPA 1994	124[7]	1[7]	USGS 1998
Mercury	1.4 µg/L	0.77 µg/L	USEPA 1999					
Nickel	470 µg/L	52 µg/L	USEPA 1999	780 µg/L	USEPA 1994	102[5]	15[2]	EC and HC 1994; CCREM 1987
Silver	3.4 µg/L		USEPA 1999					
Zinc	120 µg/L	120 µg/L	USEPA 1999	73 µg/L	USEPA 1994	51[7]	10[7]	USGS 1998
Polycyclic Aromatic Hydrocarbons (PAHs)								
Acenaphthene								
Acenaphthylene								
Anthracene								
Fluorene								
2-Methylnaphthalene								
Naphthalene								
Phenanthrene								
Benz(a)anthracene								

Table 2.6. Summary of potential targets for pore water chemistry.

Analyte	WQ Criteria		Reference	LC_{50} for *Hyalella azteca*	Reference	Invertebrates		Reference
	Acute	Chronic				Acute	Chronic	
PAHs (cont.)								
Dibenz(a,h)anthracene								
Benzo(a)pyrene								
Chrysene								
Fluoranthene								
Pyrene								
Total PAHs								
Polychlorinated Biphenyls (PCBs)								
Aroclor 1016		0.014 µg/L	USEPA 1999					
Aroclor 1221		0.014 µg/L	USEPA 1999					
Aroclor 1232		0.014 µg/L	USEPA 1999					
Aroclor 1242		0.014 µg/L	USEPA 1999					
Aroclor 1248		0.014 µg/L	USEPA 1999					
Aroclor 1254		0.014 µg/L	USEPA 1999					
Aroclor 1260		0.014 µg/L	USEPA 1999					
Total PCBs		0.014 µg/L	USEPA 1999					
Pesticides								
Chlordane	2.4 µg/L	0.0043 µg/L	USEPA 1999					
Dieldrin	0.24 µg/L	0.056 µg/L	USEPA 1999					
sum DDD								
sum DDE								
sum DDT	1.1 µg/L	0.001 µg/L	USEPA 1999					
Total DDT	1.1 µg/L	0.001 µg/L	USEPA 1999					
Endrin	0.086 µg/L	0.036 µg/L	USEPA 1999					
Heptachlor	0.52 µg/L	0.0038 µg/L	USEPA 1999					

Table 2.6. Summary of potential targets for pore water chemistry.

Analyte	WQ Criteria		Reference	LC_{50} for *Hyalella azteca*	Reference	Invertebrates		Reference
	Acute	Chronic				Acute	Chronic	
Pesticides (cont.)								
Heptachlor epoxide	0.52 µg/L	0.0038 µg/L	USEPA 1999					
Lindane (gamma-BHC)	0.95 µg/L		USEPA 1999					
Others								
Phenol								
Ammonia (total)		[a]	USEPA 1999					

[a] Temperature and pH dependent

Table 3.1. Rating of selection criteria for freshwater sediment toxicity testing organisms (ASTM 2001a; USEPA 2000a).

Criterion	*Hyalella azteca*	*Diporeia* spp.	*Chironomus tentans*	*Chironomus riparius*	*Lumbriculus variegatus*	*Tubifex tubifex*	*Hexagenia* spp.	Molluscs	*Daphnia* spp. *and Ceriodaphnia* spp.
Relative sensitivity toxicity database	+	-	+	-	+	-	-	-	-
Round-robin studies conducted	+	-	+	-	-	-	-	-	-
Contact with sediment	+	+	+	+	+	+	+	+	-
Laboratory culture	+	-	+	+	+	+	-	-	+
Taxonomic identification	+	+/-	+/-	+/-	+	+	+	+	+
Ecological importance	+	+	+	+	+	+	+	+	+
Geographical distribution	+	+/-	+	+	+	+	+	+	+/-
Sediment physicochemical tolerance	+	+	+/-	+	+	+	-	+	NA
Response confirmed with benthos populations	+	+	+	+	+	+	+	-	+
Peer reviewed	+	+	+	+	+	+	+	-	+/-
Endpoints monitored	S,G,M	S,B,A	S,G,E	S,G,E	B,S	S,R	S,G	B	S,G,R
Overall Assessment	10+	5+	8+	7+	9+	8+	5+	5+	4+

"+" or "-" rating indicates a positive or negative attribute; NA = not applicable.
S - survival; G = growth; M = maturation; E = emergence; B = bioaccumulation; R = reproduction.

Table 3.2. Summary of standard methods for conducting whole-sediment toxicity or sediment bioaccumulation tests with freshwater invertebrates.

Species	Common Name	Duration of Exposure (days)	Primary Endpoints	Standard Method	Matching Chemistry and Toxicity Data[b]
Hyalella azteca	Amphipod	10 to 14	Survival and growth	ASTM (2001b); Environment Canada (1997a); USEPA (2000b)	673 and 670
Hyalella azteca	Amphipod	28 to 42	Survival, growth, and reproduction	ASTM (2001b); USEPA (2000b)	165 and 160
Diporeia spp.	Amphipods	28	Survival and bioaccumulation	ASTM (2001b)	Not reported
Chironomus tentans	Midge	10 to 14	Survival, emergence, and growth	ASTM (2001b); Environment Canada (1997b); USEPA (2000b)	556 and 557
Chironomus tentans[a]	Midge	20 to 60	Survival, growth, emergence, and reproduction	ASTM (2001b); USEPA (2000b)	Not reported
Chironomus riparius	Midge	10 to 14	Survival and growth	Environment Canada (1997b)	76 and 81
Chironomus riparius[a]	Midge	30	Survival, growth, and emergence	ASTM (2001b)	Not reported
Daphnia magna or *Ceriodaphnia dubia*	Cladocerans	7	Survival and reproduction	ASTM (2001b)	8
Hexagenia spp.	Mayflies	21	Survival and growth	ASTM (2001b)	112
Tubifex tubifex	Oligochaete	28	Survival and reproduction	ASTM (2001b)	Not reported
Lumbriculus variegatus[a]	Oligochaete	28	Bioaccumulation	ASTM (2001d); USEPA (2000b)	Not reported

[a] OECD is currently developing standard methods for conducting sediment tests with these species (tests with *Chironomus yoshimatsui* are also being developed).

[b] Number of samples with matching sediment chemistry and toxicity in a national database described in USEPA (2000b).

Table 3.3. Advantages and disadvantages of laboratory sediment toxicity tests (ASTM 2001a; USEPA 2000a).

Advantages	Disadvantages
• Measure bioavailable fraction of contaminant(s).	• Sediment collection, handling, and storage can alter sediment toxicity.
• Provide a direct measure of benthic effects, assuming no field adaptation or amelioration of effects.	• Spiked sediment may not be representative of field-contaminated sediment.
• Limited special equipment is required.	• Natural geochemical characteristics of sediment may affect the response of test organisms.
• Methods are rapid and inexpensive.	• Indigenous animals may be present in field-collected sediments.
• Legal and scientific precedence exist for use; ASTM standard guides are available.	• Route of exposure may be uncertain and data generated in sediment toxicity tests may be difficult to interpret if factors controlling the bioavailability of contaminants in sediment are unknown.
• Measure unique information relative to chemical analyses or benthic community analyses.	• Tests applied to field samples may not discriminate effects of individual chemicals.
• Tests with spiked chemicals provide data on cause-effect relationships.	• Few comparisons have been made of methods or species.
• Sediment-toxicity tests can be applied to all COPCs.	• Only a few chronic methods for measuring sublethal effects have been developed or extensively evaluated.
• Tests applied to field samples reflect cumulative effects of contaminants and contaminant interactions.	• Laboratory tests have inherent limitations in predicting ecological effects.
• Toxicity tests are amenable to confirmation with natural benthos populations.	• Tests do not directly address human health effects.

Table 3.4. Test conditions for conducting 28- to 42-day sediment toxicity test with *Hyalella azteca* (ASTM 2001a; USEPA 2000a).

Parameter	Conditions
Test type	Whole-sediment toxicity test with renewal of overlying water.
Temperature	23 ± 1°C.
Light quality	Wide-spectrum fluorescent lights.
Illuminance	About 100 to 1000 lux.
Photoperiod	16L:8D.
Test chamber	300-mL high-form lipless beaker.
Sediment volume	100 Ml.
Overlying water volume	175 mL in the sediment exposure from Day 0 to Day 28 (175 to 275 mL in the water-only exposure from Day 28 to Day 42).
Renewal of overlying water	2 volume additions/d; continuous or intermittent (e.g., one volume addition every 12 h).
Age of organisms	7- to 8-d old at the start of the test.
Number of organisms/chamber	10
Number of replicate chambers/treatment	12 (4 for 28-day survival and growth and 8 for 35- and 42-day survival, growth, and reproduction). Reproduction is more variable than growth or survival; hence, more replicates might be needed to establish statistical differences among treatments.

Table 3.4. Test conditions for conducting 28- to 42-day sediment toxicity test with *Hyalella azteca* (ASTM 2001a; USEPA 2000a).

Parameter	Conditions
Feeding	YCT food, fed 1.0 mL (1800 mg/L stock) daily to each test chamber.
Aeration	None, unless dissolved oxygen in overlying water drops below 2.5 mg/L.
Overlying water	Culture water, well water, surface water or site water. Use of reconstituted water is not recommended.
Test chamber cleaning	If screens become clogged during a test; gently brush the outside of the screen.
Overlying water quality	Hardness, alkalinity, conductivity, and ammonia at the beginning and end of a sediment exposure (Day 0 and 28). Temperature daily. Conductivity weekly. Dissolved oxygen (DO) and pH three times/week. Concentrations of DO should be measured more often if DO drops more than 1 mg/L since the previous measurement.
Test duration	42 days.
Endpoints	28-day survival and growth; 35- and 42-day survival, growth, reproduction, and number of adult males and females on Day 42.
Test acceptability	Minimum mean control survival of 80% on Day 28. Additional performance-based criteria specifications are outlined in Table 3.5 and in round-robin.

Table 3.5. Test acceptability requirements for a 42-day sediment toxicity test with *Hyalella azteca* (ASTM 2001a; USEPA 2000a).

It is recommended for conducting the 42-day test with *Hyalella azteca* that the following performance criteria be met:	• Age of *Hyalella azteca* at the start of the test should be 7- to 8-day old. Starting a test with substantially younger or older organisms may compromise the reproductive endpoint. • Average survival of *Hyalella azteca* in the control sediment on Day 28 should be greater than or equal to 80%. • Laboratories participating in round-robin testing (ASTM 2001a, USEPA 2000a) reported after 28-day sediment exposures in a control sediment (West Bearskin), survival >80% for >88% of the laboratories; length >3.2 mm/individual for >71% of the laboratories; and dry weight >0.15 mg/individual for 66% of the laboratories. Reproduction from Day 28 to Day 42 was >2 young/female for 71% of the laboratories participating in the round-robin testing. Reproduction was more variable within and among laboratories; hence, more replicates might be needed to establish statistical differences among treatments with this endpoint. • Hardness, alkalinity, and ammonia in the overlying water typically should not vary by more than 50% during the sediment exposure, and dissolved oxygen should be maintained above 2.5 mg/L in the overlying water.
Performance-based criteria for culturing *Hyalella azteca* include the following:	• It may be desirable for laboratories to periodically perform 96-h water-only reference-toxicity tests to assess the sensitivity of culture organisms. Data from these reference toxicity tests could be used to assess genetic strain or life-stage sensitivity of test organisms to select chemicals. • Laboratories should track parental survival in the cultures and record this information using control charts if known-age cultures are maintained. Records should also be kept on the frequency of restarting cultures and the age of brood organisms. • Laboratories should record the following water-quality characteristics of the cultures at least quarterly: pH, hardness, alkalinity, and ammonia. Dissolved oxygen in the cultures should be measured weekly. Temperature in the cultures should be recorded daily. If static cultures are used, it may be desirable to measure water quality more frequently.

Table 3.5. Test acceptability requirements for a 42-day sediment toxicity test with *Hyalella azteca* (ASTM 2001a; USEPA 2000a).

Performance-based criteria (cont.)	• Laboratories should characterize and monitor background contamination and nutrient quality of food if problems are observed in culturing or testing organisms. • Physiological measurements such as lipid content might provide useful information regarding the health of the cultures.
Additional requirements:	• All organisms in a test must be from the same source. • Storage of sediments collected from the field should follow guidance outlined in ASTM (2000a) and in USEPA (2000a). • All test chambers (and compartments) should be identical and should contain the same amount of sediment and overlying water. • Negative-control sediment and appropriate solvent controls must be included in a test. The concentration of solvent used must not adversely affect test organisms. • Test organisms must be cultured and tested at 23°C (±1 °C). • The mean of the daily test temperature must be within ± 1°C of 23°C. The instantaneous temperature must always be within ±3°C of 23°C. • Natural physico-chemical characteristics of test sediment collected from the field should be within the tolerance limits of the test organisms.

Table 3.6. Uncertainty associated with sediment phases used in laboratory toxicity tests (Ingersoll *et al.* 1997).

	Whole Sediment: Benthos	Whole Sediment: Pelagic	Organic Extracts	Suspended Solids	Elutriates	Pore Water
Precision	1	1	1	3	1	1
Ecological relevance	1	2	3	2	3	2
Causality: Link	3	3	3	3	3	3
Causality: Source	1	2	3*	3	3	2
Sensitivity	1	2	3	3	3	2
Interference	2*	2	3	3	3	2*
Standardization	1	2	3	3	1	2
Discrimination	1*	1*	1*	1*	1*	1*
Bioavailability	1*	1*	3	1*	3	1*
Field validation	1*	2*	3	3*	3	3*

Ranking Code: 1 = low uncertainty (good); 3 = high (bad); * = lack of knowledge.

Table 3.7. Uncertainty associated with endpoints measured in laboratory toxicity tests with sediment (Ingersoll *et al.* 1997).

	Survival	Growth	Reproduction	Behavior	Life Tables	Development	Biomarkers
Precision	1	1*	2*	1*	3*	1	2*
Ecological relevance	1	2*	1*	2*	1*	2*	3*
Causality: Link	3	3	3	3	3	3	2
Causality: Source	1	2*	2*	2*	3*	1	2
Sensitivity	1	2	1	2	2*	2	1*
Interference	1*	2*	3*	2*	3*	2*	3
Standardization	1	2	2	1	3	2*	3
Discrimination	2	1	1	2	2*	1	2*
Bioavailability	1	1	1	1	1	1	1
Field validation	1	2*	2*	1	3*	2	3

Ranking Code: 1 = low uncertainty (good); 3 = high (bad); * = lack of knowledge.

Table 4.1. Uncertainty associated with benthic community assessments (Ingersoll *et al.* 1997).

	Individual	Population	Structure	Function
Precision	1	1	2	3*
Ecological relevance	3	2	1	3*
Causality: Contamination	2*	2*	2*	3*
Causality: Source	2*	3*	3*	3*
Sensitivity	1*	1	2	3*
Interference	2*	3*	3*	3*
Standardization	3*	1	1	3*
Discrimination	2	1	1	3*
Bioavailability	2*	NA	NA	3*
Field validation	3*	1	1	3*

Ranking Code: 1 = low uncertainty (good); 3 = high (bad); * = lack of knowledge.
NA = not applicable.

Table 4.2. Advantages and disadvantages of benthic invertebrate community structure data.

Advantages	Disadvantages
• Provides information that is directly relevant for assessing the status of the benthic community. • Procedures are available to facilitate defensible sampling program design. • Resultant data are socially- and ecologically-relevant. • Limited special equipment is required to support assessments.	• The distribution and abundance of benthic invertebrates can be influenced by non-contaminant related factors (e.g., TOC, grain size). • Large numbers of samples are needed to address the inherent variability of benthic community metrics. • Standard methods for collecting and processing samples are not available. • Identification of organisms to species can be difficult. • Benthic community data can not be used alone to determine the cause of any effects that are observed. • There is little agreement on which metrics are the most relevant for use in benthic community assessments.

Table 5.1. Selection criteria for sediment bioaccumulation test organisms (ASTM 2001d; USEPA 2000a).

Criterion	*Lumbriculus variegatus*	Molluscs	Midges	Mayflies	Amphipods	Cladocerans	Fish
Laboratory culture	+	-	+	-	+	+	+
Known chemical exposure	+	-	+	+/-	+	+	+
Adequate tissue mass	+/-	+	-	+	-	-	+
Low sensitivity to contaminants	+	+	-	-	-	-	+/-
Feeding not required during testing	+	+	-	+	-	-	+
Realistic exposure	+	+/-	+	+	+	-	-
Sediment physico-chemical tolerance	+	?	+/-	-	+	NA	NA
Response confirmed with benthic populations	+	?	?	?	+	?	-
Overall assessment	7+	3+	3+	3+	5+	2+	4+

"+" or "-" rating indicates a positive or negative attribute.
NA = not applicable; ? = unknown.

Table 5.2. Advantages and disadvantages of tissue chemistry data.

Advantages	Disadvantages
• Provides direct information for determining the presence/absence of COPCs in tissues.	• Can not be used to evaluate effects on ecological receptors directly.
• Standard methods are available for most COPCs.	• Generation of high quality data can require substantial sample volumes, which is difficult to obtain for small organisms or for areas that have depauperate benthic communities.
• Procedures are available for evaluating the reliability of the data (i.e., accuracy and precision).	• Effective interpretation of the data is dependent on the availability of appropriate benchmarks.
• Benchmarks (i.e., TRGs) are available for many COPCs for evaluating the potential for biological effects.	• The use of inappropriate methods (e.g., with high detection limits) can limit the utility of the resultant data.
• Can be used to identify the COPCs that are causing or substantially contributing to adverse effects.	• Interferences with the analysis of specific analytes can influence the utility of the data (i.e., by resulting in high detection limits).

Table 5.3. Recommended test conditions for conducting a 28-day sediment bioaccumulation test with *Lumbriculus variegatus* (ASTM 2001d; USEPA 2000a).

Parameter	Conditions
Test type	Whole-sediment bioaccumulation test with renewal of overlying water.
Temperature	23°C.
Light quality	Wide-spectrum fluorescent lights.
Illuminance	About 100 to 1000 lx.
Photoperiod	16L:8D.
Test chamber	4 to 6-L aquaria with stainless steel screens or glass standpipes.
Sediment volume	1 L or more depending on TOC.
Overlying water volume	1 L or more depending on TOC.
Renewal of overlying water	2 volume additions/day; continuous or intermittent (for example, one volume addition every 12 h).
Age of test organisms	Adults.
Loading of organisms in chamber	Ratio of TOC in sediment to organism dry weight should be no less than about 50:1; minimum of 1 g/replicate; preferably 5 g/replicate.
Number of replicate chambers/treatment	Depends on the objective of the test. Five replicates are recommended for routine testing.

Table 5.3. Recommended test conditions for conducting a 28-day sediment bioaccumulation test with *Lumbriculus variegatus* (ASTM 2001d; USEPA 2000a).

Parameter	Conditions
Feeding	None.
Aeration	None, unless dissolved oxygen in overlying water drops below 2.5 mg/L.
Overlying water	Culture water, well water, surface water, site water, or reconstituted water.
Test chamber cleaning	If screens become clogged during the test, gently brush the *outside* of the screen.
Overlying water quality	Hardness, alkalinity, conductivity, pH, and ammonia at the beginning and end of a test temperature and dissolved oxygen daily.
Test duration	28 days.
Endpoint	Bioaccumulation.
Test acceptability	Performance-based criteria specifications outlined in Table 5.4.

Table 5.4. Test acceptability requirements for a 28-day sediment bioaccumulation test with the oligochaete, *Lumbriculus variegatus* (ASTM 2001d; USEPA 2000a).

It is recommended for conducting a 28-day test with *Lumbriculus variegatus* that the following performance criteria are met:	• Numbers of *Lumbriculus variegatus* in a 4-day toxicity screening test should not be reduced significantly in the test sediment relative to the control sediment. • Test organisms should burrow into test sediment. Avoidance of the test sediment by *Lumbriculus variegatus* may decrease bioaccumulation. • The hardness, alkalinity, pH, and ammonia of overlying water within a treatment typically should not vary by more than 50 % during the test and dissolved oxygen should be maintained above 2.5 mg/L in the overlying water.
Performance-based criteria for culturing *Lumbriculus variegatus* include the following:	• It may be desirable for laboratories to perform periodically 96-hour water-only reference toxicity tests to assess the sensitivity of culture organisms. Data from these reference toxicity tests could be used to assess genetic strain or life-stage sensitivity of test organisms to select chemicals. • Laboratories should monitor the frequency with which the population is doubling in the culture (the number of organisms) and record this information using control charts (the doubling rate would need to be estimated on a subset of animals from a mass culture). Records also should be kept on the frequency of restarting cultures. If static cultures are used, it may be desirable to measure water quality more frequently. • Food used to culture organisms should be analyzed before the start of a test for compounds to be evaluated in the bioaccumulation test. • Laboratories should record the following water quality characteristics of the cultures at least quarterly and the day before the start of a sediment test: pH, hardness, alkalinity, and ammonia. Dissolved oxygen in the cultures should be measured weekly. Temperatures of the cultures should be recorded daily. • Laboratories should characterize and monitor the background contamination and nutrient quality of food if problems are observed in culturing or testing organisms.

Table 5.4. Test acceptability requirements for a 28-day sediment bioaccumulation test with the oligochaete, *Lumbriculus variegatus* (ASTM 2001d; USEPA 2000a).

Performance-based criteria (cont.)	• Physiological measurements such as lipid content might provide useful information regarding the health of the cultures.
Additional requirements:	• All organisms in a test must be from the same source.
	• Storage of sediment collected from the field should follow guidance outlined in ASTM (2001).
	• All test chambers (and compartments) should be identical and should contain the same amount of sediment and overlying water.
	• Negative-control sediment or appropriate solvent controls, must be included in a test. The concentration of solvent used must not affect test organisms adversely.
	• Culture and test temperatures must be the same. Acclimation of test organisms to the test water is not required.
	• The daily mean test temperature must be within ±1°C of the desired temperature. The instantaneous temperature must always be within ±3°C of the desired temperature.
	• Natural physicochemical characteristics of test sediment collected from the field should be within the tolerance limits of the test organisms.

Table 5.5. Uncertainty associated with bioaccumulation assessments (Ingersoll *et al.* 1997).

	Laboratory	Field	Food Web	Models
Precision	1	2	3	3
Ecological relevance: Protection of ecology	3	3	3	3
Ecological relevance: Protection of human health	1	1	1	1
Causality: Source identification	1	3	3	1
Causality: Sensitivity (detection limit)	1	2	3*	3*
Interferences	2	2*	NA	NA
Standardization	1	1	2	2
Discrimination	1	1	1	1
Bioavailability	1	1	1	1
Field validation	2*	1	2*	2*

Ranking Code: 1 = low uncertainty (good); 3 = high (bad); * = lack of knowledge.
NA = not applicable.

Table 6.1. Methods for evaluating the effects of exposure to COPCs in fish (from Schmitt *et al.* 2000).

Method	Description	Tissue(s) Examined	Sensitivity	Reference
Histopathology	Microscopic examination for the presence of lesions; can provide early indication of chemical exposure	Liver, gill, gonads, spleen, and kidney	Overall organism health and contaminants	Hinton *et al.* 1992; Hinton 1993; Goodbred *et al.* 1997
Ethoxyresorufin-*O*-deethylase (EROD) activity	Enzyme induction by planar hydrocarbons	Liver	PCBs, PAHs, dioxins, and furans	Pohl and Fouts 1980; Kennedy and Jones 1994
Lysozyme activity	A disease resistance factor that can be suppressed in the presence of contaminants	Blood plasma	Overall organism health	Blazer *et al.* 1994a
Macrophage aggregate analysis	Macrophages are important in the immune system, serving as a first line of defense for the organism and as an antigen processing cell	Spleen, hemopoetic kidney, and liver	Multiple contaminants including PAHs and metals	Blazer *et al.* 1994a; Blazer *et al.* 1997
H4IIE bioassay	A screening tool to determine the presence of certain classes of planar halogenated compounds	Whole fish (composites)	PCBs, dioxins, furans, and PAHs	Tillitt *et al.* 1991
Vitellogenin	A precursor of egg yolk, normally synthesized in the liver of female fish	Blood plasma	Endocrine modulating compounds	Folmar *et al.* 1996
Sex Steroids (estradiol and testosterone)	Determine reproductive health and status	Blood plasma	Endocrine modulating compounds	Guillete *et al.* 1994; Goodbred *et al.* 1997

Table 6.1. Methods for evaluating the effects of exposure to COPCs in fish (from Schmitt *et al.* 2000).

Method	Description	Tissue(s) Examined	Sensitivity	Reference
Chemical analyses	Organochlorine chemical residues and elemental contaminants	Whole fish (composites)	Specific analytes	Schmitt *et al.* 1999
Somatic indices	The relative mass of some organs is often indicative of chemical exposure	Gonads, spleen, liver	Overall organism health	Grady *et al.* 1992
Stable N isotopes (^{14}N and ^{15}N)	The ratio of (^{15}N to ^{14}N) (d^{15}N) increases with trophic position and sewage pollution	Whole fish (composites)	Trophic position, nitrogen sources	Grady *et al.* 1996
Necropsy-based fish health assessment	Visual assessment of external/internal anomalies (e.g., lesions, parasites, tumors), which may indicate contaminant-related stress	All	Overall organism health	Goede 1988; 1996; Adams 1990; Adams *et al.* 1993

Table 6.2. Methodological uncertainty associated with fish health and fish community assessments.

	Fish Health	Fish Community
Precision	1	2
Ecological relevance	2	1
Causality	1	3
Sensitivity	2	2
Interference	3	3
Standardization	2	2
Discrimination	1	3
Bioavailability	1	3
Field validation	2	2

Ranking Code: 1 = low uncertainty (good); 3 = high (bad).

Table 7.1. Contingency table for assessing impacts of contaminated sediments on aquatic life based on three separate indicators of sediment quality (sediment quality triad adapted from Chapman 1992 and Canfield *et al.* 1996).

Possible Outcome	Sediment Chemistry	Toxicity Test	Benthic Community	Possible Conclusions
1	+	+	+	Impact highly likely: Contaminant-induced degradation of sediment-dwelling organisms evident.
2	-	-	-	Impact highly unlikely: Contaminant-induced degradation of sediment dwelling organisms not evident.
3	+	-	-	Impact unlikely: Contaminants unavailable to sediment-dwelling organisms.
4	-	+	-	Impacts possible: Unmeasured contaminants or conditions exist that have the potential to cause degradation.
5	-	-	+	Impacts unlikely: No degradation of sediment-dwelling organisms in the field apparent relative to sediment contamination; physical factors may be influencing benthic community.
6	+	+	-	Impact likely: Toxic chemicals probably stressing the system.
7	-	+	+	Impact likely: Unmeasured toxic chemicals are probably contributing to the toxicity.
8	+	-	+	Impact likely: Sediment-dwelling organisms degraded by toxic chemicals, but toxicity tests not sensitive to chemicals present.

+ = Indicator classified as affected; as determined based on comparison to the established target.

- = Indicator not classified as affected; as determined based on comparison to the established target.

Table 7.2. Contingency table for assessing impacts of contaminated sediments on aquatic life based on four separate indicators of sediment quality.

Possible Outcome	Sediment Chemistry	Toxicity Test	Benthic Community	Tissue Chemistry	Possible Conclusions
1	+	+	+	+	Contaminant-induced impacts on sediment-dwelling organisms and higher trophic levels are likely to be observed; elevated levels of sediment-associated contaminants are likely contributing to sediment toxicity and benthic community impairment; and, bioaccumulation of sediment-associated contaminants has the potential to adversely affect aquatic-dependent wildlife and/or human health.
2	-	-	-	+	Contaminant-induced impacts on higher trophic levels are likely to be observed; adverse effects on sediment-dwelling organisms are unlikely to be observed; and, bioaccumulation of sediment-associated contaminants has the potential to adversely affect aquatic-dependent wildlife and/or human health.
3	+	-	-	+	Contaminant-induced impacts on higher trophic levels are likely to be observed; the bioavailability of sediment-associated contaminants is likely to be limited; and, bioaccumulation of sediment-associated contaminants has the potential to adversely affect aquatic-dependent wildlife and/or human health.
4	-	+	-	+	Contaminant-induced impacts on higher trophic levels are likely to be observed; unmeasured factors (e.g., physical factors or contaminants) are likely to be contributing to sediment toxicity; and, bioaccumulation of sediment-associated contaminants has the potential to adversely affect aquatic-dependent wildlife and/or human health.
5	-	-	+	+	Contaminant-induced impacts on sediment-dwelling organisms and higher trophic levels are likely to be observed; adverse effects on sediment-dwelling organisms are likely due to physical factors and/or unmeasured chemicals are stressing benthos and toxicity tests are not sensitive enough to detect effects; and, bioaccumulation of sediment-associated contaminants has the potential to adversely affect aquatic-dependent wildlife and/or human health.

Table 7.2. Contingency table for assessing impacts of contaminated sediments on aquatic life based on four separate indicators of sediment quality.

Possible Outcome	Sediment Chemistry	Toxicity Test	Benthic Community	Tissue Chemistry	Possible Conclusions
6	+	+	-	+	Contaminant-induced impacts on sediment-dwelling organisms and higher trophic levels are likely to be observed; high variability in the benthic community metrics may be masking contaminant-related effects; and, bioaccumulation of sediment-associated contaminants has the potential to adversely affect aquatic-dependent wildlife and/or human health.
7	-	+	+	+	Contaminant-induced impacts on sediment-dwelling organisms and higher trophic levels are likely to be observed; unmeasured contaminants are likely contributing to sediment toxicity and benthic impairment; and, bioaccumulation of sediment-associated contaminants has the potential to adversely affect aquatic-dependent wildlife and/or human health.
8	+	-	+	+	Contaminant-induced impacts on sediment-dwelling organisms and higher trophic levels are likely to be observed; toxicity tests are not sensitive enough to detect adverse effects; and, bioaccumulation of sediment-associated contaminants has the potential to adversely affect aquatic-dependent wildlife and/or human health.
9	+	+	+	-	Contaminant-induced impacts on sediment-dwelling organisms are likely to be observed; elevated levels of sediment-associated contaminants are likely contributing to sediment toxicity and benthic community impairment; and, bioaccumulation of sediment-associated contaminants is unlikely to be adversely affect aquatic-dependent wildlife and/or human health.
10	-	-	-	-	Contaminant-induced impacts are unlikely to be observed; sediment-associated contaminants are unlikely to adversely affect sediment-dwelling organisms; and, bioaccumulation of sediment-associated contaminants is unlikely to adversely affect aquatic-dependent wildlife and/or human health.

Table 7.2. Contingency table for assessing impacts of contaminated sediments on aquatic life based on four separate indicators of sediment quality.

Possible Outcome	Sediment Chemistry	Toxicity Test	Benthic Community	Tissue Chemistry	Possible Conclusions
11	+	-	-	-	Contaminant-induced impacts are unlikely to be observed; the bioavailability of sediment-associated contaminants is likely to be limited; and, bioaccumulation of sediment-associated contaminants is unlikely to adversely affect aquatic-dependent wildlife and/or human health.
12	-	+	-	-	Contaminant-induced impacts are unlikely to be observed, based on the COPCs that were evaluated; Unmeasured factors (e.g., physical factors or contaminants) are likely to be contributing to sediment toxicity; and, bioaccumulation of sediment-associated contaminants is unlikely to adversely affect aquatic-dependent wildlife and/or human health.
13	-	-	+	-	Contaminant-induced impacts on sediment-dwelling organisms are unlikely to be observed, based on the COPCs that were evaluated; adverse effects on sediment-dwelling organisms are likely due to physical factors and/or unmeasured chemicals are stressing benthos and toxicity tests are not sensitive enough to detect effects; and, bioaccumulation of sediment-associated contaminants is unlikely to adversely affect aquatic-dependent wildlife and/or human health.
14	+	+	-	-	Contaminant-induced impacts on sediment-dwelling organisms are likely to be observed; high variability in the benthic community metrics may be masking contaminant-related effects; and, bioaccumulation of sediment-associated contaminants is unlikely to adversely affect aquatic-dependent wildlife and/or human health.
15	-	+	+	-	Contaminant-induced impacts on sediment-dwelling organisms are likely to be observed, based on the COPCs that were evaluated; unmeasured contaminants are likely contributing to sediment toxicity and benthic impairment; and, bioaccumulation of sediment-associated contaminants is unlikely to adversely affect aquatic-dependent wildlife and/or human health.

Table 7.2. Contingency table for assessing impacts of contaminated sediments on aquatic life based on four separate indicators of sediment quality.

Possible Outcome	Sediment Chemistry	Toxicity Test	Benthic Community	Tissue Chemistry	Possible Conclusions
16	+	-	+	-	Contaminant-induced impacts on sediment-dwelling organisms are likely to be observed; toxicity tests are not sensitive enough to detect adverse effects; and, bioaccumulation of sediment-associated contaminants is unlikely to adversely affect aquatic-dependent wildlife and/or human health.

+ = Indicator classified as affected; as determined based on comparison to the established target.

- = Indicator not classified as affected; as determined based on comparison to the established target.

Table 7.3. Contingency table for assessing impacts of contaminated sediments on aquatic life based on two separate indicators of sediment quality.

Possible Outcome	Sediment Chemistry	Sediment Toxicity	Possible Conclusions
1	+	+	Impact likely: Contaminant-induced degradation of sediment-dwelling organisms evident.
2	-	-	Impact unlikely: Contaminant-induced degradation of sediment-dwelling organisms not evident.
3	+	-	Impact unlikely: Chemicals not readily available to sediment-dwelling organisms.
4	-	+	Impact likely: Observed effects likely due to unmeasured contaminants or physical factors.

+ = Indicator classified as affected; as determined based on comparison to the established target.

- = Indicator not classified as affected; as determined based on comparison to the established target.

Table 7.4. Contingency table for assessing impacts of contaminated sediments on wildlife based on three separate indicators of sediment quality.

Possible Outcome	Sediment Chemistry	Fish Community	Tissue Chemistry	Possible Conclusions
1	+	+	+	Impact likely: Contaminant-induced effects on wildlife in the field and bioaccumulation evident.
2	-	-	-	Impact unlikely: Contaminant-induced effects on wildlife in the field not evident; limited bioaccumulation.
3	+	-	-	Impact unlikely: Contaminants unavailable to wildlife in the field.
4	-	+	-	Impact unlikely: Effects on wildlife in the field probably not due to sediment contamination; limited bioaccumulation.
5	-	-	+	Impact unlikely: No degradation of wildlife in the field apparent relative to sediment contamination; tissue residues due to exposure from other media and/or sites.
6	+	+	-	Impact likely: Contaminant induced effects on wildlife in the field; bioaccumulative substances not contributing to effects.
7	-	+	+	Impact unlikely: Effects on wildlife in the field probably not due to contaminated sediment; bioaccumulation may be occurring due to exposure at other sites.
8	+	-	+	Impact likely: Contaminants not toxic to wildlife, but bioaccumulation is occurring.

+ = Indicator classified as affected; as determined based on comparison to the established target.
- = Indicator not classified as affected; as determined based on comparison to the established target.

Table 7.5. Contingency table for assessing impacts of contaminated sediments on human health based on two separate indicators of sediment quality.

Possible Outcome	Sediment Chemistry	Tissue Chemistry	Possible Conclusions
1	+	+	Impact likely: Elevated sediment chemistry and tissue residues resulting in potential adverse dietary affects on human health.
2	-	-	Impact unlikely: Sediment chemistry and tissue residues low, with limited potential of adverse dietary affects on human health.
3	+	-	Impact possible: Sediment chemistry elevated to level that may result in potential adverse dietary affects on human health, but organisms sampled for tissue chemistry may not be exposed to sediments at the site or contaminants are not readily available.
4	-	+	Impact possible: Elevated tissue residues resulting in potential adverse dietary affects on human health, but organisms are probably not exposed to sediments at the site.

+ = Indicator classified as affected; as determined based on comparison to the established target.
- = Indicator not classified as affected; as determined based on comparison to the established target.

Table A1.1. Incidence of toxicity predicted in laboratory toxicity tests using mean PEC-quotients (USEPA 2000b).

Test Species/Duration	Incidence of Toxicity (%) by Mean PEC-Q			
	<0.1	0.1 - <0.5	0.5 - <1.0	>1.0
Hyalella azteca , 10 to 14-day	18%	16%	37%	54%
Hyalella azteca , 28 to 42-day	10%	13%	56%	97%
Chironomus spp ., 10 to 14-day	20%	17%	43%	52%

Figures

Figure 2.1. Recommended procedure for assessing sediment chemistry data.

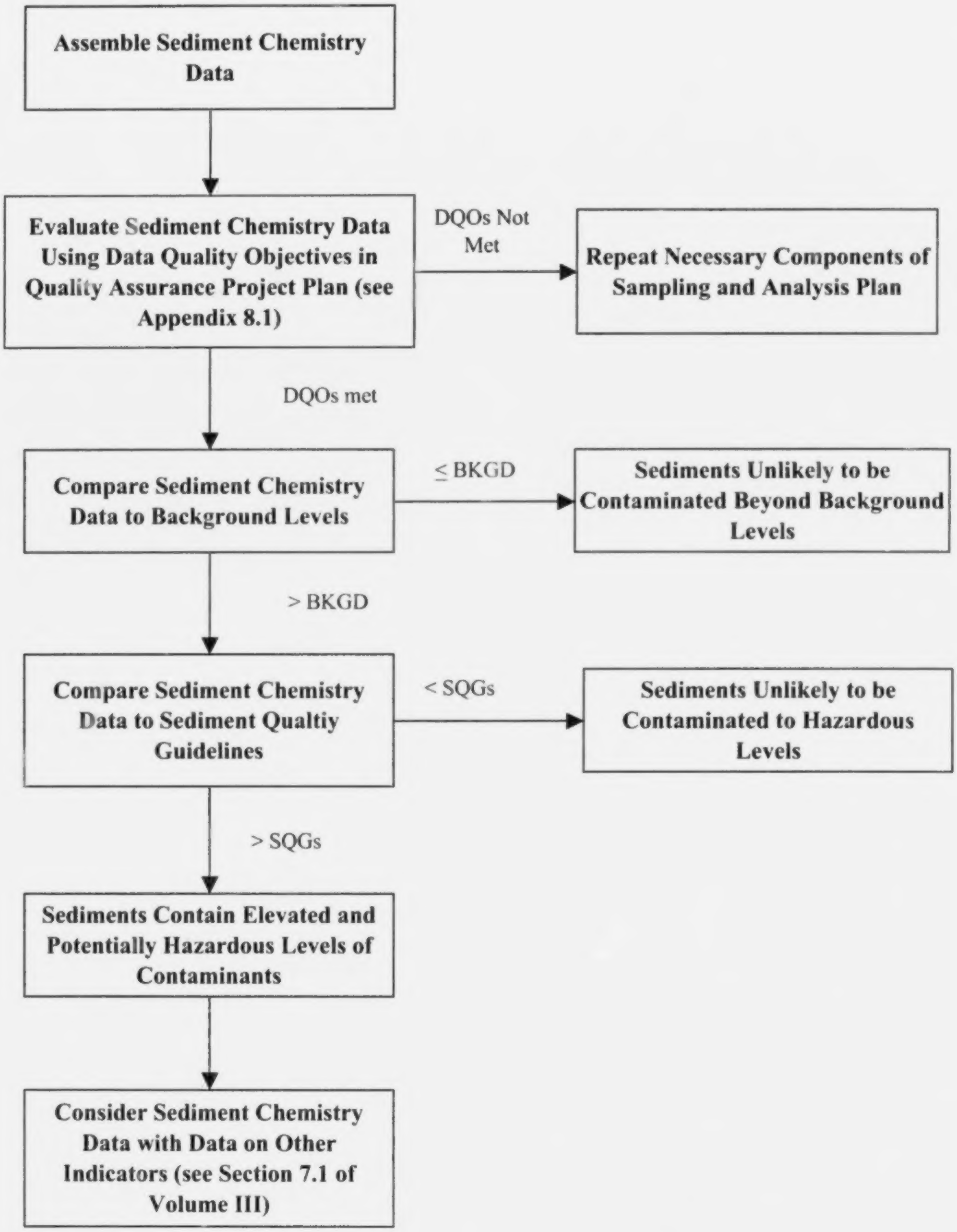

Assemble Sediment Chemistry Data
Evaluate Sediment Chemistry Data Using Data Quality Objectives in Quality Assurance Project Plan (see Appendix 8.1)
DQOs Not Met
Repeat Necessary Components of Sampling and Analysis Plan
DQOs met
Compare Sediment Chemistry Data to Background Levels
≤ BKGD
Sediments Unlikely to be Contaminated Beyond Background Levels
> BKGD
Compare Sediment Chemistry Data to Sediment Qualtiy Guidelines
< SQGs
Sediments Unlikely to be Contaminated to Hazardous Levels
> SQGs
Sediments Contain Elevated and Potentially Hazardous Levels of Contaminants
Consider Sediment Chemistry Data with Data on Other Indicators (see Section 7.1 of Volume III)

Figure 3.1. Relationship between mean PEC quotients and the incidence of toxicity in freshwater toxicity tests (USEPA 2000b).

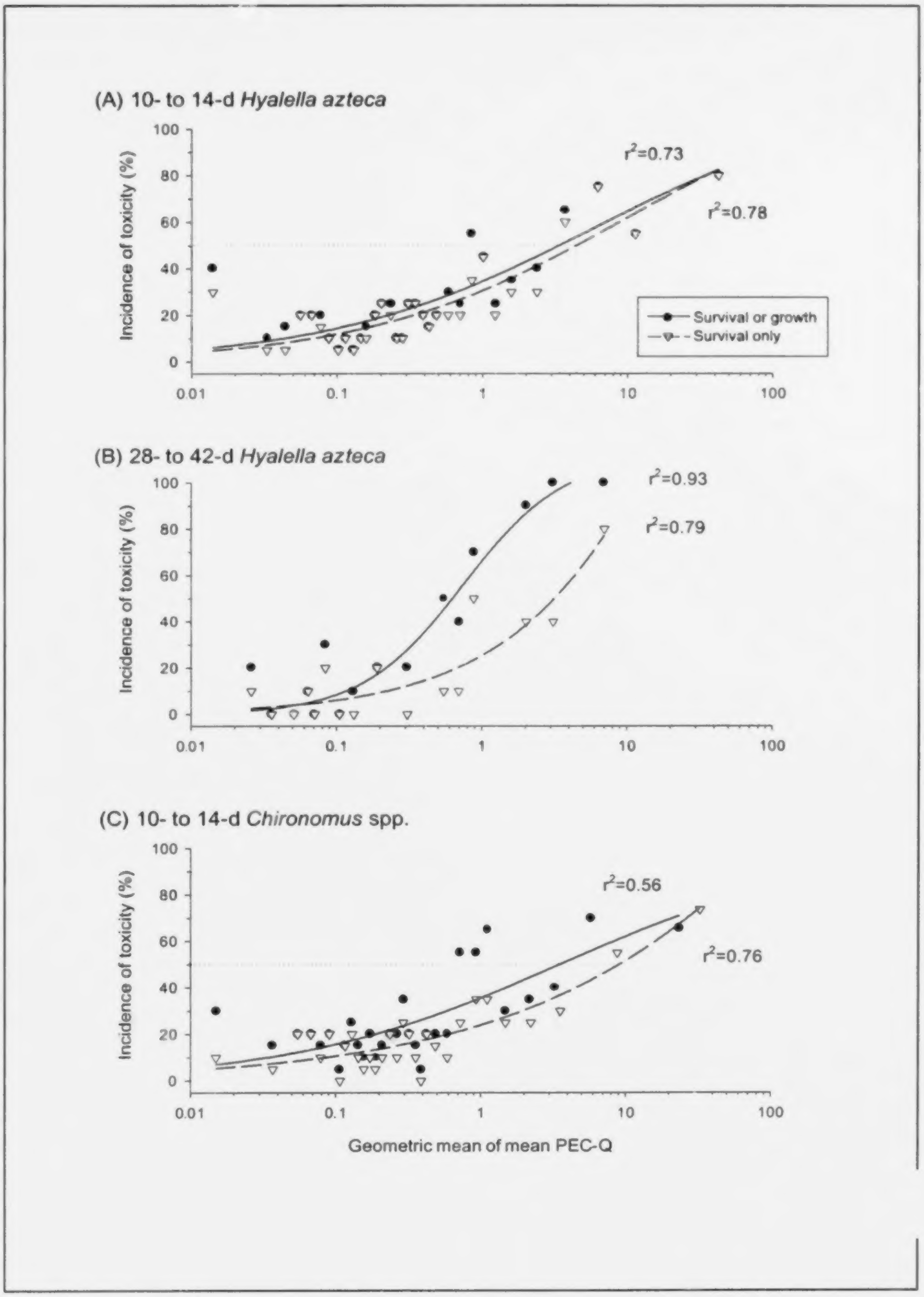
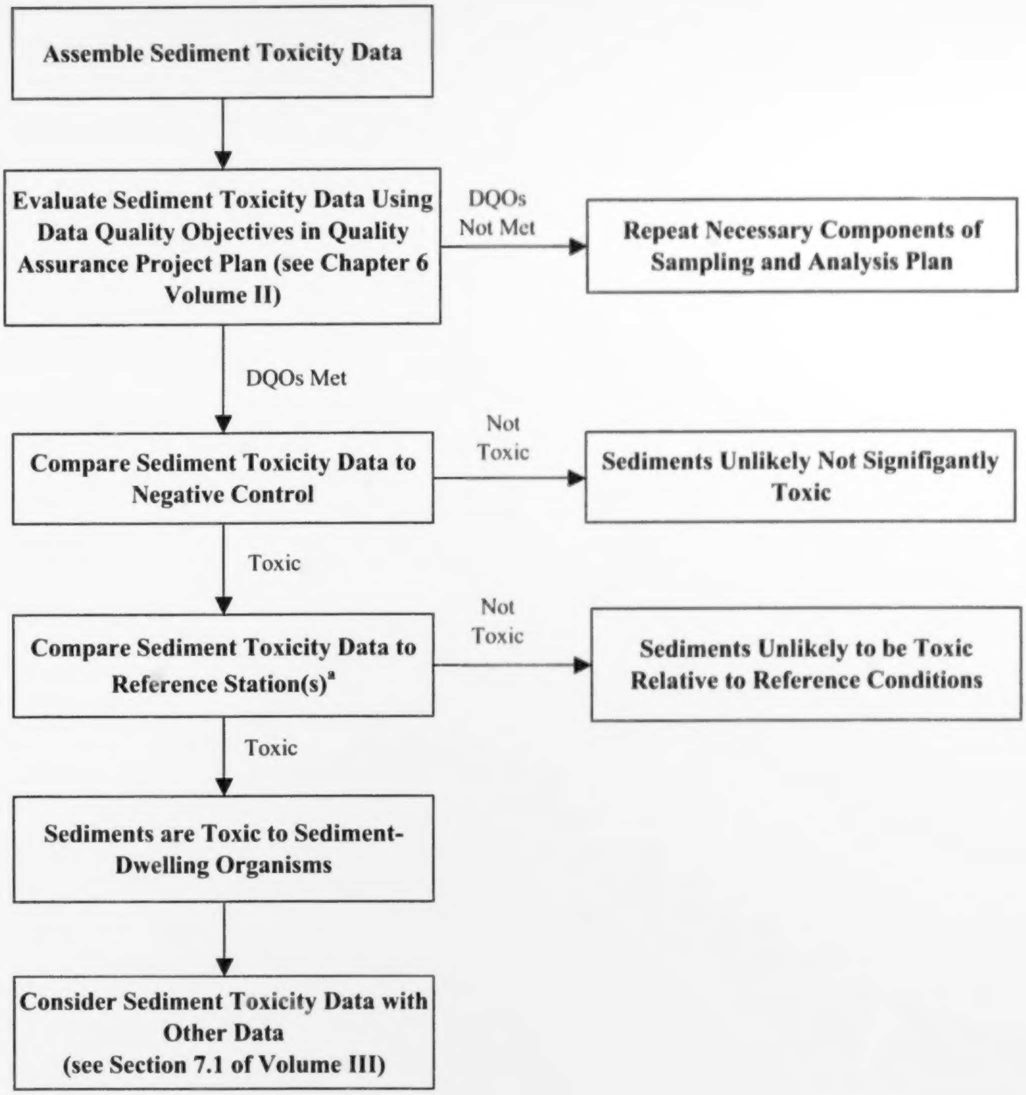
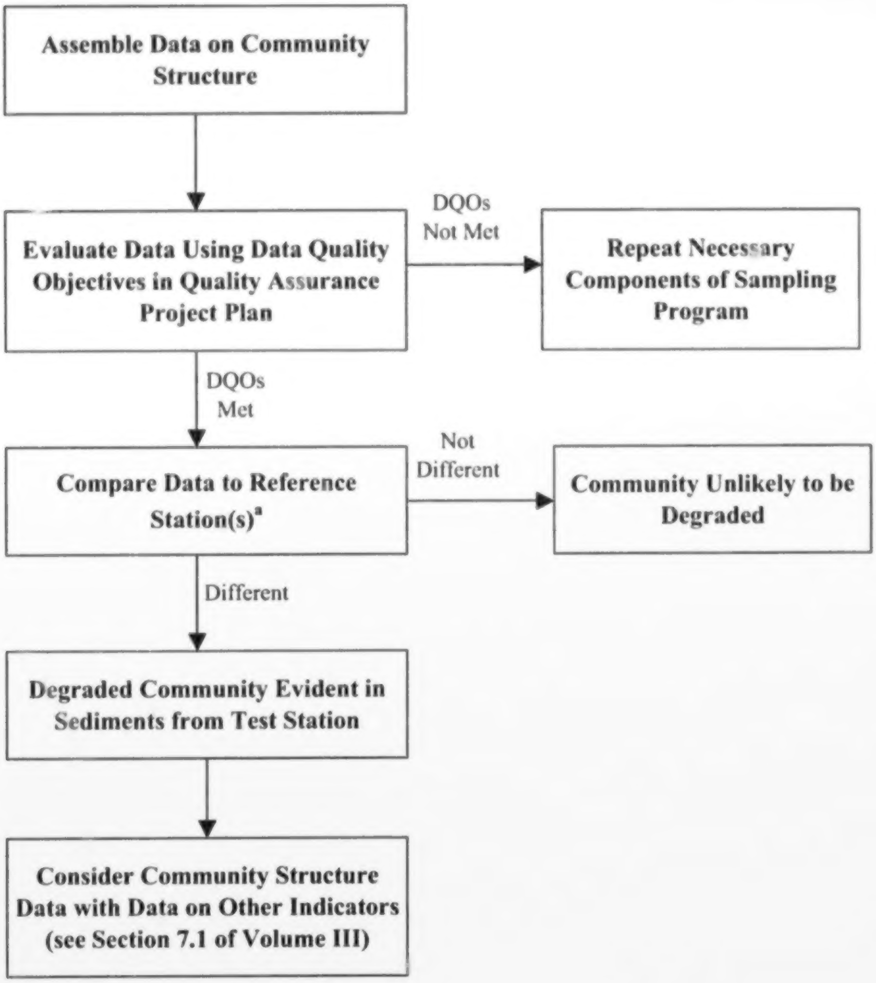


Figure 3.2. Recommended procedure for assessing sediment toxicity data.


Assemble Sediment Toxicity Data
Evaluate Sediment Toxicity Data Using Data Quality Objectives in Quality Assurance Project Plan (see Chapter 6 Volume II)
DQOs Not Met
Repeat Necessary Components of Sampling and Analysis Plan
DQOs Met
Compare Sediment Toxicity Data to Negative Control
Not Toxic
Sediments Unlikely Not Signifigantly Toxic
Toxic
Compare Sediment Toxicity Data to Reference Station(s)[a]
Not Toxic
Sediments Unlikely to be Toxic Relative to Reference Conditions
Toxic
Sediments are Toxic to Sediment-Dwelling Organisms
Consider Sediment Toxicity Data with Other Data (see Section 7.1 of Volume III)

[a] Comparison to reference sites is only appropriate if reference sites have been well charactized and satisfy criteria for negative controls (i.e., response in reference sediments should not be significantly different from that in negative controls).

Figure 4.1. Recommended procedure for assessing benthic invertebrate or fish community structure.


Assemble Data on Community Structure
Evaluate Data Using Data Quality Objectives in Quality Assurance Project Plan
DQOs Not Met
Repeat Necessary Components of Sampling Program
DQOs Met
Compare Data to Reference Station(s)a
Not Different
Community Unlikely to be Degraded
Different
Degraded Community Evident in Sediments from Test Station
Consider Community Structure Data with Data on Other Indicators (see Section 7.1 of Volume III)

[a] Comparison to reference sites is only appropriate if reference sites have been well charactized and satisfy criteria for negative controls (i.e., response in reference sediments should not be significantly different from that in negative controls).

Figure 5.1. Recommended procedure for assessing tissue chemistry data.

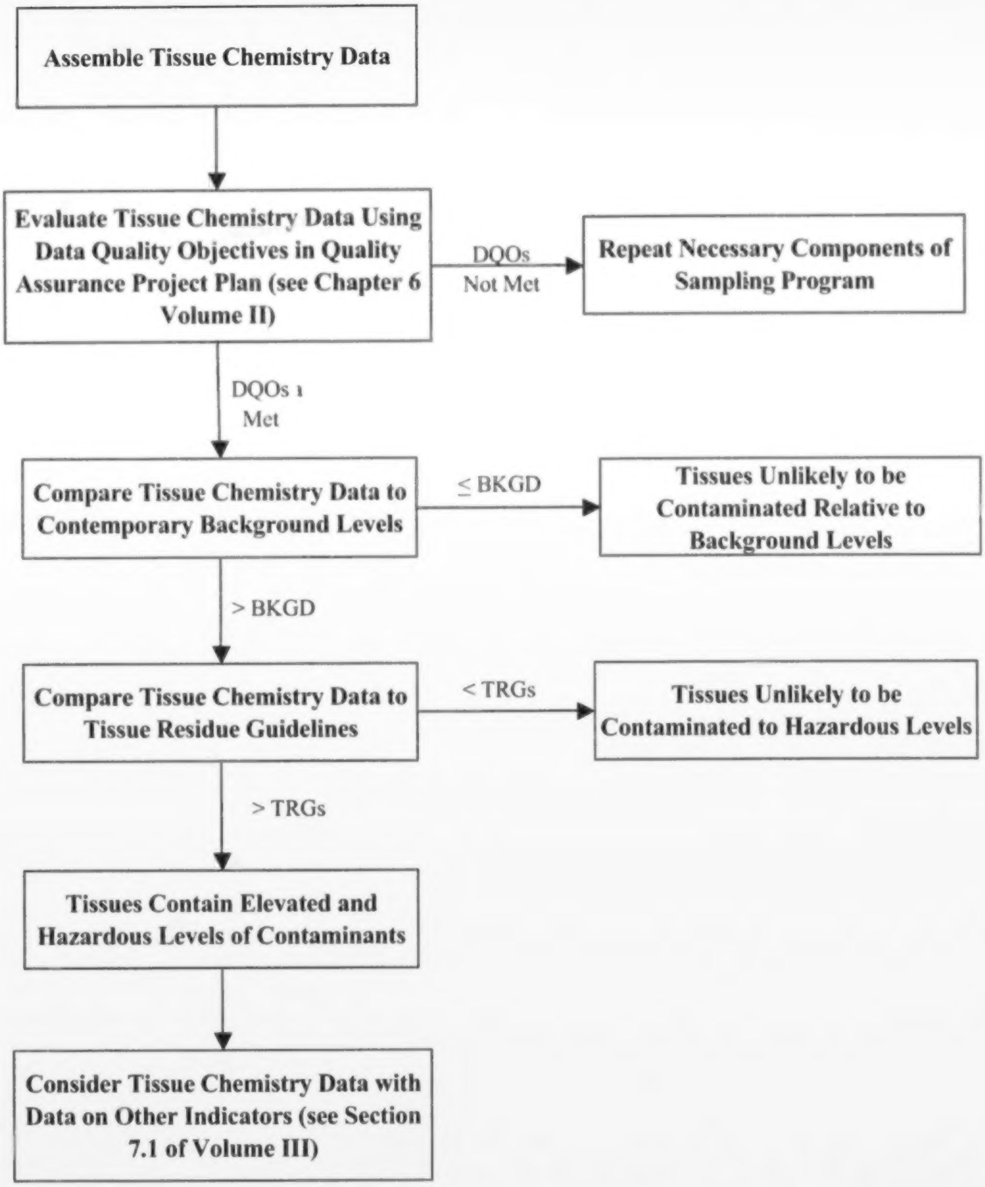

Assemble Tissue Chemistry Data
Evaluate Tissue Chemistry Data Using Data Quality Objectives in Quality Assurance Project Plan (see Chapter 6 Volume II)
DQOs Not Met
Repeat Necessary Components of Sampling Program
DQOs Met
Compare Tissue Chemistry Data to Contemporary Background Levels
≤ BKGD
Tissues Unlikely to be Contaminated Relative to Background Levels
> BKGD
Compare Tissue Chemistry Data to Tissue Residue Guidelines
< TRGs
Tissues Unlikely to be Contaminated to Hazardous Levels
> TRGs
Tissues Contain Elevated and Hazardous Levels of Contaminants
Consider Tissue Chemistry Data with Data on Other Indicators (see Section 7.1 of Volume III)

Figure 6.1. Recommended procedure for evaluating fish health data.

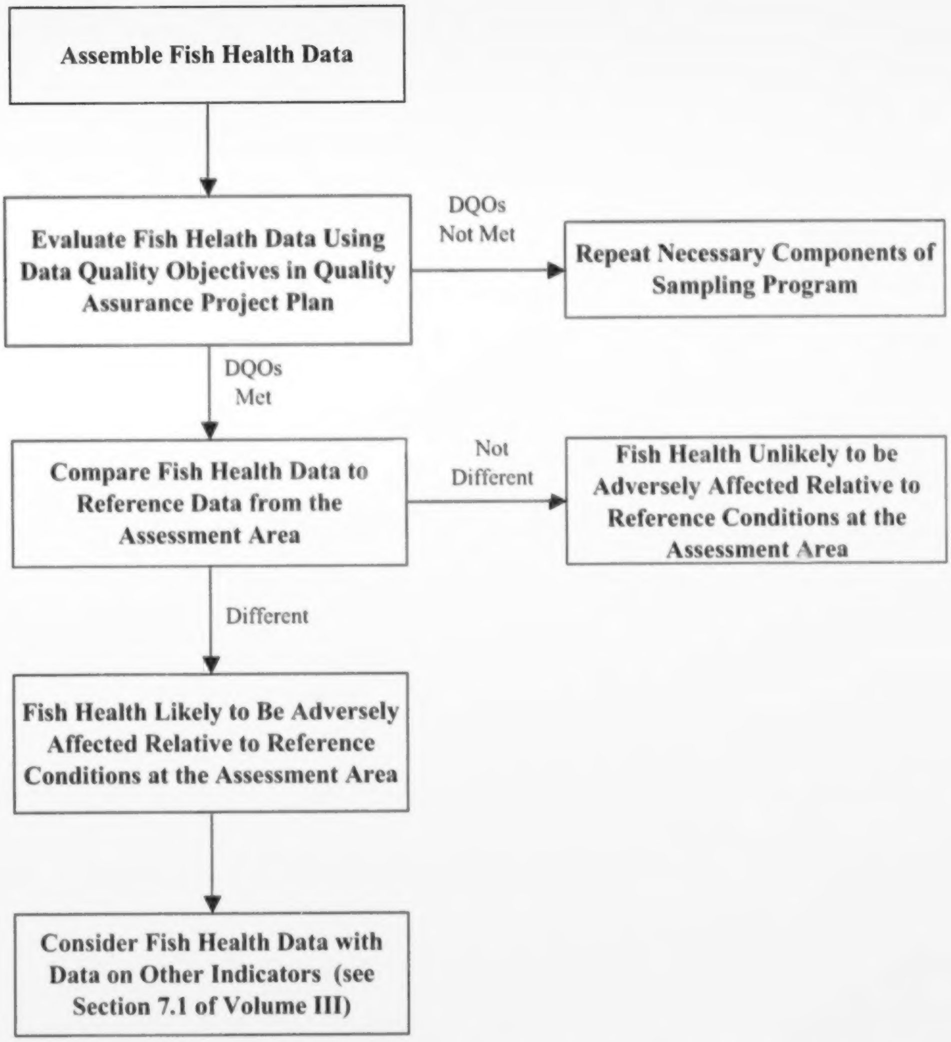

Assemble Fish Health Data
Evaluate Fish Helath Data Using Data Quality Objectives in Quality Assurance Project Plan
DQOs Not Met
Repeat Necessary Components of Sampling Program
DQOs Met
Compare Fish Health Data to Reference Data from the Assessment Area
Not Different
Fish Health Unlikely to be Adversely Affected Relative to Reference Conditions at the Assessment Area
Different
Fish Health Likely to Be Adversely Affected Relative to Reference Conditions at the Assessment Area
Consider Fish Health Data with Data on Other Indicators (see Section 7.1 of Volume III)

Figure 7.1. The relationship between the mean PEC quotient and the response of *Hyalella azteca* in the 10-day tests (as percent survival) or the response in the Microtox® solid-phase sediment toxicity test (as the EC_{50} expressed as a toxicity reference index). Sediment samples were collected from the Grand Calumet River and Indiana Harbor Canal located in northwestern Indiana (Ingersoll *et al.* 2001b).

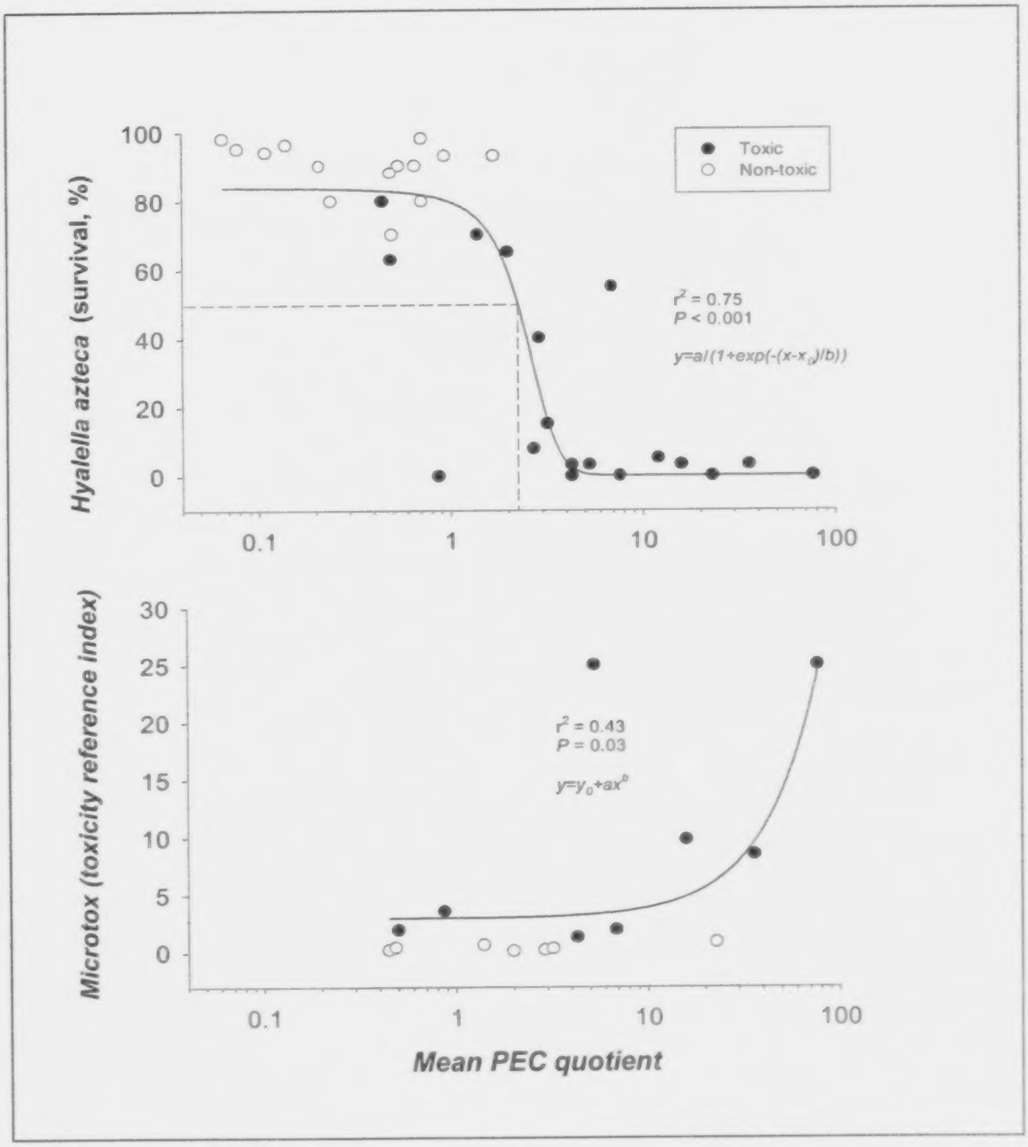


Figure 7.2. The relationship between the molar concentration of simultaneously extracted metals to acid volatile sulfide (SEM-AVS) and toxic units of metals in the sediment samples. Toxicity of samples was determined using 10-day whole-sediment tests with *Hyalella azteca* (Ingersoll *et al.* 2001b).

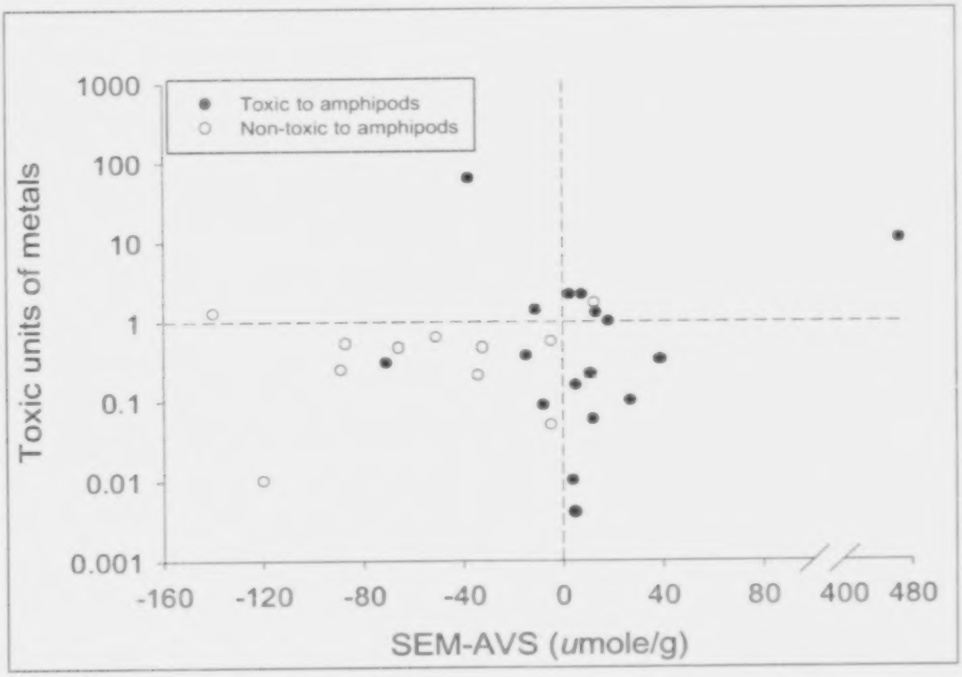


Figure 7.3 Tri-axial graphs of sediment quality triad data (Canfield *et al.* 1994; C = chemistry, T = toxicity, and B = benthic community; see Section 7.1 of Volume II for description of metrics).

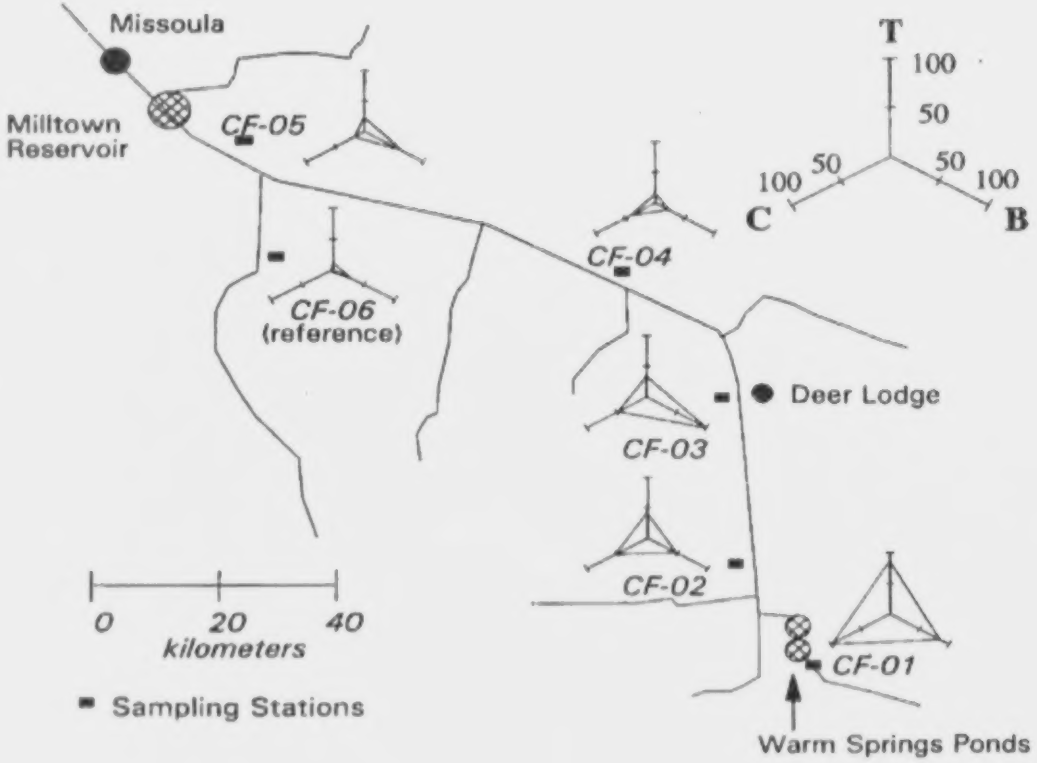

Missoula
Milltown Reservoir
CF-05
CF-06 (reference)
CF-04
CF-03
CF-02
CF-01
Deer Lodge
Warm Springs Ponds
T
100
50
C
100
50
B
50
100
0
20
40
kilometers
Sampling Stations

Figure A2.1. Metal/aluminum regression lines with the 95% prediction limits (from Carvalho and Schropp 2001).

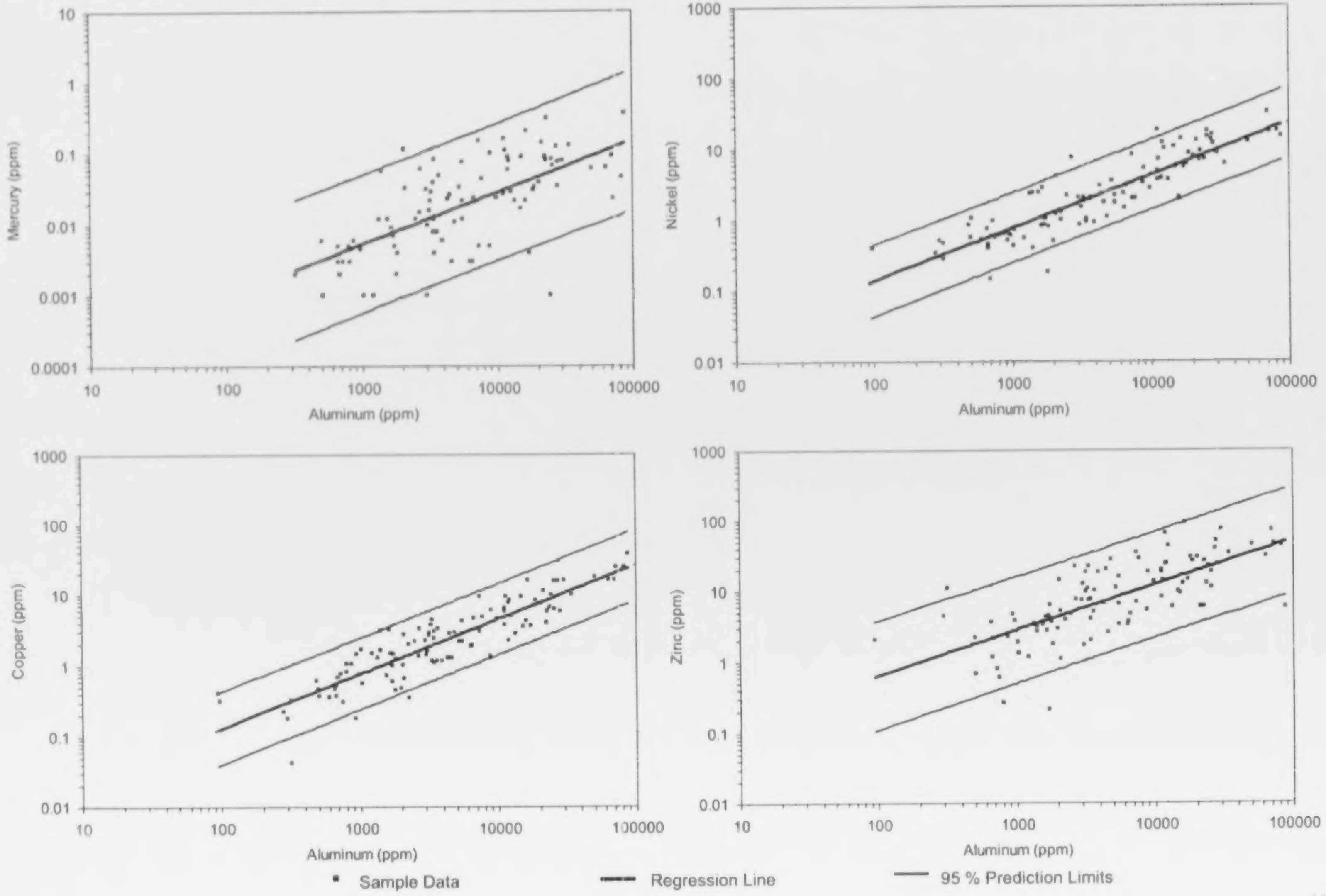


Figure A2.1. Metal/aluminum regression lines with the 95% prediction limits (from Carvalho and Schropp 2001).

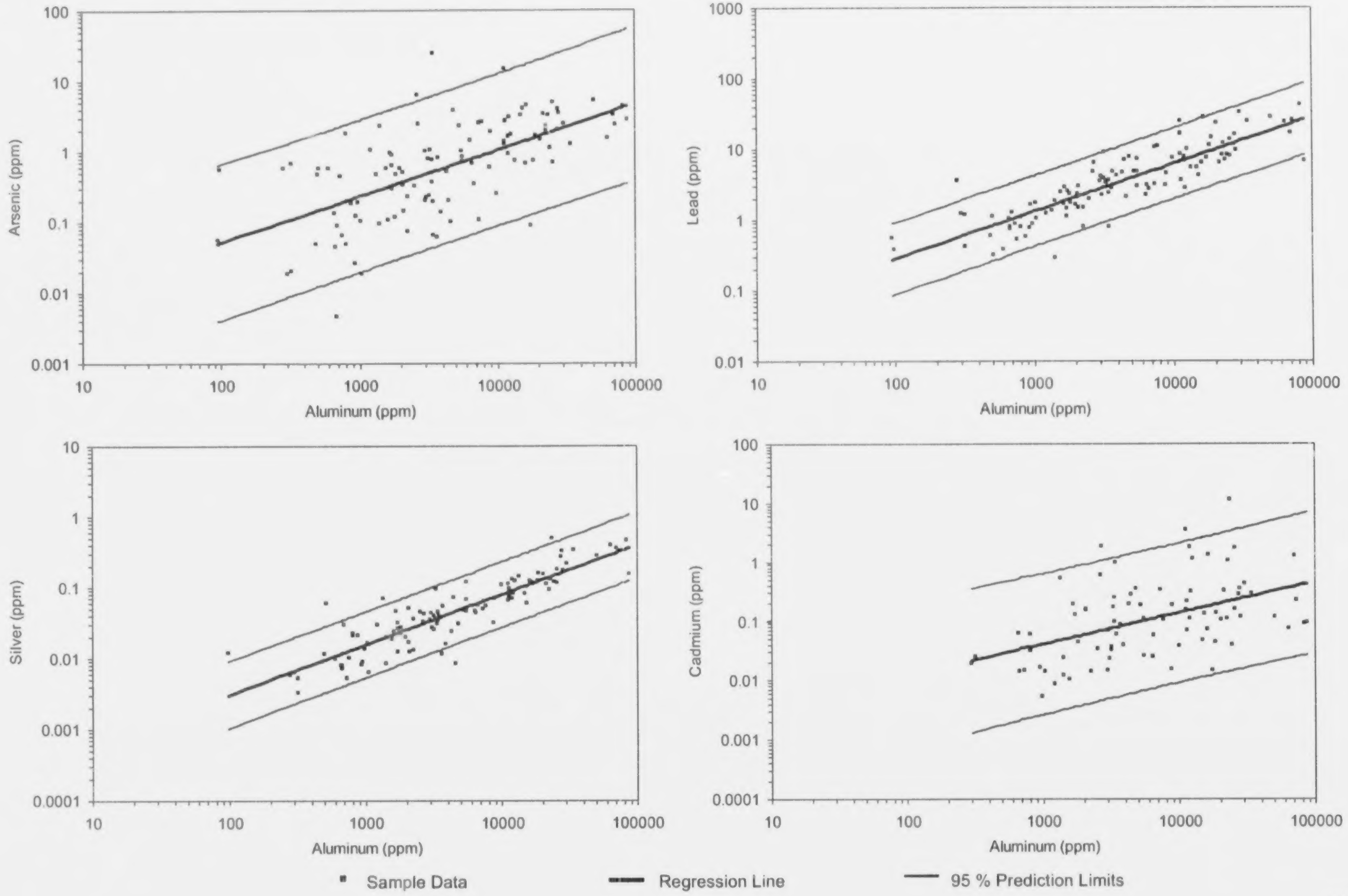


Figure A2.1. Metal/aluminum regression lines with the 95% prediction limits (from Carvalho and Schropp 2001).